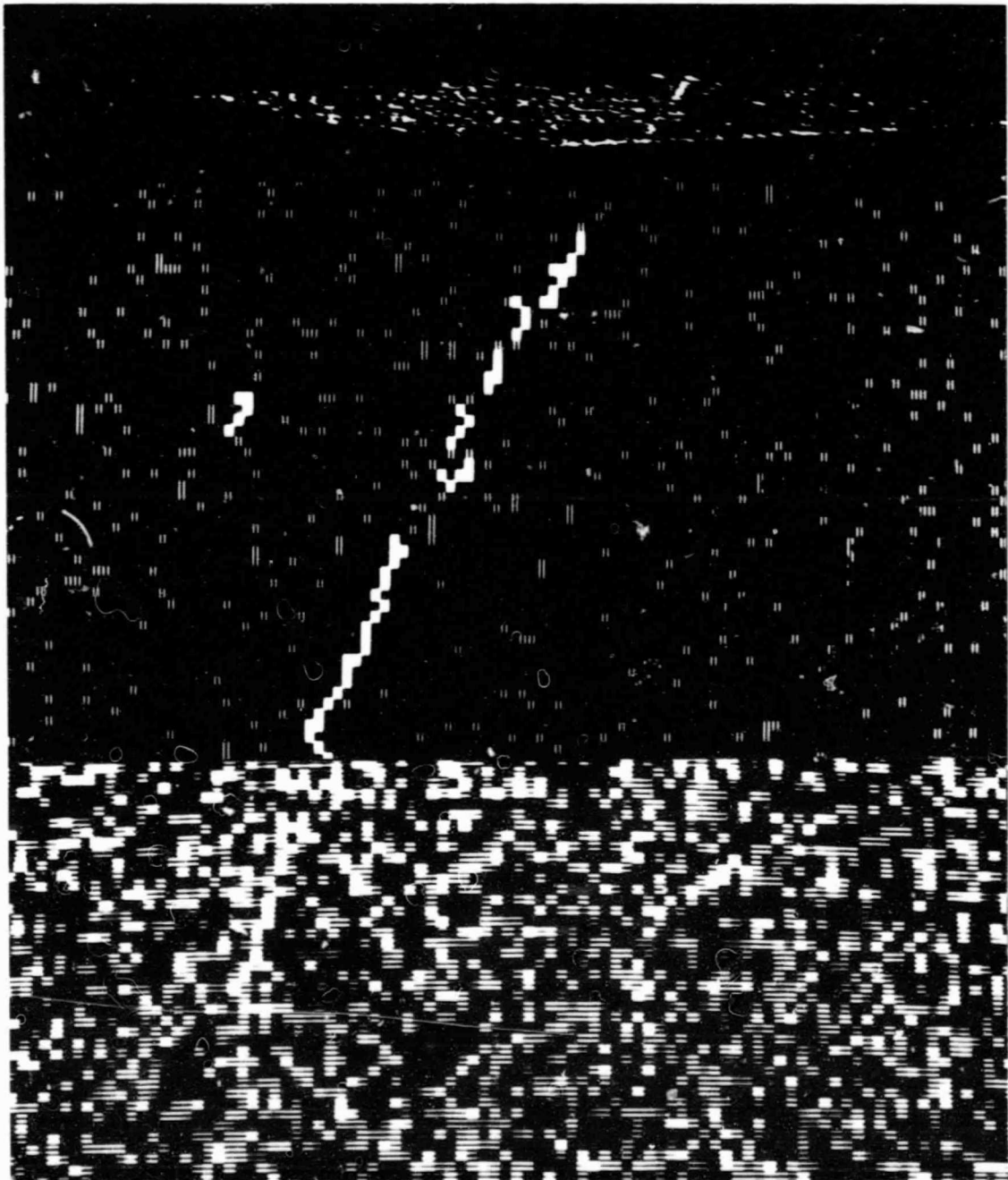


# PROJECT OASIS:

## THE DESIGN OF A SIGNAL DETECTOR FOR

## The Search for Extraterrestrial Intelligence



**PROJECT OASIS:**

**THE DESIGN OF A  
SIGNAL DETECTOR FOR**

**The Search for Extraterrestrial Intelligence**

**A technical report from the 1979 Summer Study at  
NASA-Ames Research Center, Moffett Field, California**

**Study Directors**

**Timothy Healy, University of Santa Clara  
Mark Stull, NASA Ames Research Center**

**Editor**

**Steven Lord, University of Massachusetts**

**Associate Editors**

**Robert Dixon, Ohio State University  
Timothy Healy, University of Santa Clara**

**University of Santa Clara  
Santa Clara, California  
1981**



## List of Participants

John Bahng  
Department of Physics and Astronomy  
Northwestern University

Jack Baird  
Department of Psychology  
Dartmouth College

Madelaine Bates  
Mathematics Department  
Bronx Community College

Tyler Blake  
Department of Psychology  
California State University

Steve Choquette  
University of Santa Clara

Kent Culler  
Department of Computer Science  
University of California, Berkeley

John Day  
Infomatics

Stanley Deans  
Physics Department  
University of South Florida

Rose Dishman  
United States International University

Bob Dixon  
Radio Observatory Department  
Ohio State University

Jeremy Dunning  
Geology Department  
Indiana University

Bill Evans  
University of Santa Clara

Steven Lord  
Department of Physics and Astronomy  
University of Massachusetts

Frank Maloney  
Department of Astronomy  
Villanova University

Scott McClelland  
NASA Ames Research Center

Alex Nauda  
Electrical Engineering Department  
Bucknell University

Norm Rhodine  
Department of Electrical Engineering  
University of Wyoming

Gerhard Ritter  
Department of Mathematics  
University of Florida

Bob A. Rubin  
Mathematics Department  
Whittier College

Bob H. Rubin  
Physics Department  
California State University

Jim Schmandle  
University of Santa Clara

Gene Sogliero  
Mathematics Department  
Trinity College

Charles Spiteri  
Department of Electrical Technology  
Queensborough Community College

Gerry Zeitlin  
Space Sciences Laboratory  
University of California, Berkeley

"There is no use trying," she said: "One can't believe impossible things." "I daresay you haven't had much practice," said the Queen. "When I was your age, I always did it for half-an-hour a day. Why, sometimes I've believed as many as six impossible things before breakfast."  
--Lewis Carroll

## Preface

For at least a decade it has been recognized that any truly effective search of the radio frequency band for extraterrestrial signals would require a signal detector system of unprecedented resolution and sensitivity. Unfortunately, all searches to date have relied upon rather conventional detection equipment, and thus have not been able to probe deeply into our stellar neighborhood.

A great impetus for the search for extraterrestrial intelligence (SETI) effort came in 1973 in the form of a NASA-Ames Summer Study design of a large scale search system. This system, called Project Cyclops, gains sensitivity through linking a great number of telescopes of moderate size. The study demonstrated the feasibility of a full-scale SETI program to begin in the near future.

In addition to Project Cyclops, the rapid advance of digital electronics in the 70's also led to increased interest in SETI. For the first time, through the development of innovative FFT algorithms and fast microprocessors, an 8-million-channel spectrum analyzer (MCSA) was designed to meet the needs of a SETI program. The MCSA puts out a very large database at very high rates. The development of the device which follows the MCSA, called the Signal Detector, represents a major design challenge in both data processing and algorithm optimization. The design of this device is the subject of this report.



The study was sponsored by the American Society for Engineering Education (ASEE) and the National Aeronautics and Space Administration (NASA). The host institutions were the University of Santa Clara and the NASA-Ames Research Center. The 1979 Summer Study brought together 22 professional educators in the fields of physics, statistics, electrical engineering, and radio astronomy. During this summer they designed an integrated system of a new and original type which can be built with today's technology.

As the study commenced the group was presented with two very difficult problems. They were: "How does one process 1/2-terabits ( $10^{12}$  bits) of data in 1000 seconds," and "How does one detect a completely unspecified signal with acceptable sensitivity?"

Fortunately, the study team rose to the challenge, and an intense period of learning, groping, and imagining followed. Proposed solutions emerged, each with its small camp of supporters. Adding fuel to ensuing debates was the fact that for an algorithm to be acceptable it was required to be both sensitive and feasible (implementable). This requirement brought together hardware specialists and mathematical theorists. Algorithm makers conferred with data handlers. Outside specialists were frequently brought in to determine if a product could truly be produced. All candidate algorithms, such as Pattern Recognition, Analysis of Variance, and Coherence Measure were subjected to a makeshift scientific forum, where the participants learned to speak the language of signal-to-noise ratios and bit rates. Concurrent with this debate was the search for technologies commensurate to the massive quantities of processing required. New and exotic technologies (including optical laser storage, super computers, multiprocessing computer arrays) were explored. Again concurrent with these studies was one by the

psychologists of the group regarding the human brain as the most sophisticated "algorithm processor" of all. It was a period of intense interdisciplinary learning with statisticians developing integrated circuits and physicists studying human response to stimuli, and with each person periodically staring into noise-dominated display screens attempting to discern a hidden signal.

This period of open scientific query was ultimately channelled into a period of determined engineering design and layout, where all chosen methods were incorporated and integrated. It is somewhat sad that so many promising and exciting areas of research had to be eventually constrained to definitive dry statements and diagrams within this report, but it was felt that one well-specified system was needed which served to incorporate this broad scope of ideas into one.

We wish to say a word about the use of computers by the study participants. Individuals within the group produced simulations which:

- determined optimal ways of integrating drifting carrier waves\*
- tested intelligent algorithms to follow drifting carrier waves
- determined the sensitivity of coherence measure and shift SNRs
- showed the effect of quantization on variance algorithms
- developed 3-D display techniques for human pattern recognition
- demonstrated the ability of humans to recognize patterns
- demonstrated automated pattern and cluster recognition

---

\* explained in the body of this report

Such simulations with computers and with humans aided by computers helped greatly in gaining an evaluation of beliefs. Simulation of the various components of the system is an economical and effective way of evaluating and expanding the ideas proposed in this report.

This study proposes a system that acts as a detector of intelligent signals. It combines the thinking of physicists, statisticians, electrical engineers and radio astronomers into one integrated design constructable with today's technology. While we do not think that this study is the final answer in terms of instrumentation and algorithms in the search for extraterrestrial life, we do feel that many of these proposals are rich with potential and should be exploited further.

A word about the choice of the name for the study--Project Oasis. One of the major candidate regions of the electromagnetic spectrum in which to search is the range from 1400 MHz to 1700 MHz. Near the low end of this band is the spectral line of hydrogen. Near the upper end is the line for hydroxyl radical. These two lines of the dissociation products of water gave rise to the suggestion that this narrow band of frequencies, where nature passes signals with relative ease, be called the "water hole," and indeed what better place might we search than that ancient meeting ground of so many disparate species of living creatures. It was this allusion to a water hole which suggested the word Oasis, a patch of green in a vast expanse of arid desert, with its water hole providing the source of life.

We have each come from this summer project realizing the power of collaboration and cooperation between differing disciplines and viewpoints. This study is the product of a group of individuals who worked together and reached beyond their limits to break new ground, just as man must search beyond himself towards the limits of space if he is to reach the stars.

### Acknowledgments

We wish to acknowledge with gratitude the contributions which many people made to the summer study. Their efforts helped make the study a successful and rewarding experience for all of us.

The study was sponsored by the NASA-Ames Research Center and the University of Santa Clara. NASA-Ames hosted the 10-week study. Among those who helped make our stay at Ames pleasant as well as effective were Mary Connors, John Billingham, Vera Buescher, and John Leveen.

A seminar series conducted at the start of the summer provided a valuable orientation to the group, introducing us to the nature of the problem we were to attack, and suggesting many specific ideas which were useful throughout the summer. Contributors to this excellent series included: Bernard Oliver of Hewlett Packard, Jack Welch of the University of California at Berkeley, Alan Peterson of Stanford University, Jill Tarter of the University of California at Berkeley, Ron Scudder of Hewlett Packard, Charles Seeger of the Ames SETI office, Scott Knauer of NASA-Ames, Ron Bracewell of Stanford University, Theodosios Pavlidis of the University of California at Berkeley, and Jim Adams of Stanford University.

Bernard Oliver of Hewlett Packard also arranged to lend us a computer and a graphics system which enabled us to grapple with the problem of the visual detection of faint signals in noise. This excellent research tool was often used around the clock. It allowed us to conduct some specific experiments, and to develop our perception of the detection problem.

Special thanks also go to Scott Knauer for his enthusiasm and untiring assistance in our problems of hardware implementation. His innovations

included the design and construction of an analog display for viewing 3-D images, a project which well reflects his ability to transform ideas into reality.

The secretarial needs of the study were very well met by Noreen Golden and Elnora Donato. Their cheerfulness in the face of long hours of hard work was an inspiration to the group and their good spirits helped make the summer a special experience for us all.

The University of Santa Clara hosted the editing of this report. Yvette Deboscq greatly assisted in the editing task by retyping the entire manuscript, drafting many of the figures, and assisting in numerous other editorial functions.

## TABLE OF CONTENTS

Chapter 1 - Introduction - The Search for Extraterrestrial Intelligence (SETI)...	1
1.1 - The Evidence for Life .....	3
1.2 - Where is Extraterrestrial Life to Be Found?.....	4
1.3 - How Should We Look for Extraterrestrial Life? .....	5
1.4 - Anti-Cryptography vs. Anthropocentrism .....	7
1.5 - Why Should We Look? .....	8
1.6 - How Can We Understand their Messages? .....	10
1.7 - What Is Being Done Now? .....	11
1.8 - Detection Techniques Used in Previous SETI Projects .....	15
Wideband Pulsed Signal Techniques .....	15
Narrowband Continuous Signal Techniques .....	16
1.9 - The Next Step.....	18
1.10 - The Receiving System .....	19
1.11 - Purpose of This Study .....	22
1.12 - The Oasis Signal Detection System .....	22
Chapter 2 - Signals and Noise .....	26
2.1 - Introduction .....	27
2.2 - Classifying Signals and Noise .....	27
2.3 - A Global View .....	28
2.4 - What Is Noise? .....	31
2.5 - Signal Characteristics .....	32
2.6 - Signal Type Assumptions .....	38
2.7 - Review of Signal Processing Steps .....	40
2.8 - Signal Detection Algorithms .....	40

Chapter 3 - System Algorithms .....	44
3.1 - Purpose .....	45
3.2 - Overview .....	46
Real Time Processing .....	46
Signal Recognition Procedure .....	46
Human Interface .....	48
3.3 - Generalized Coherence .....	50
Background .....	50
Polarization Matching .....	51
Polarization Coherence .....	56
Time Coherence .....	59
Generalized Coherence .....	61
3.4 - Analysis of Variance (ANOVA ) Algorithm .....	67
Background .....	67
Performance .....	72
Conclusions .....	73
3.5 - Oasis System Algorithms .....	73
Overview .....	73
Data Stream Preparation .....	74
Pulse Detector (PD) .....	76
Carrier Wave Detector (CWD) .....	77
Numerical Battery of Independent Tests (NBIT) .....	79
The Individual NBIT Tests .....	86
Pattern Recognition - The Stage II NBIT Algorithm .....	95
Chapter 4 - Oasis System Architecture .....	109
4.1 - Introduction .....	110

Purpose of the System .....	110
Architectural Strategy .....	112
Large Memories and "Second Looks" at the Data .....	113
Advantages of the Architecture Chosen .....	117
4.2 - The Real-Time Processing System .....	119
Input to the Oasis Signal Detector .....	119
Array Transposer .....	121
Advantages of the Transposition Scheme .....	125
Pulse Detector Implementation .....	126
A Hardware Implementation of the Carrier Wave Detector .....	129
Implementation .....	131
The CWD Pipeline .....	135
Implementation of the NBIT Algorithm. ....	147
The Overlap System .....	151
The Common Processor Implementation .....	151
8-Hz Pulse Detector Implementation .....	154
Total Power Implementation .....	154
Goodness of Fit Implementation .....	154
Complex Coherence Implementation .....	158
ANOVA Implementation .....	158
Cluster Detector Implementation .....	162
The Histogram Circuit .....	162
Cluster Detector Microprocessor .....	164
Conclusions Concerning Realtime Processing .....	168
4.3 - Archive System Implementation .....	168
4.4 - Operator Interface .....	169
Central Computer .....	169



Threshold Setting .....	170
Detection Handling .....	170
System Log .....	171
Self-Diagnostics .....	171
Other Functions .....	172
RFI Receiver and Catalog .....	172
4.5 - Radio Astronomy System .....	173
4.6 - Gross Attributes .....	176
Storage Capacity: 2 Terabits Online .....	176
Brief Inventory of Major Components .....	177
System's Total Size .....	177
System Cost .....	179
Chapter 5 - Human Interface .....	180
5.1 - Problem Statement .....	181
5.2 - Using Human Information Processing Capabilities in SETI .....	181
Perceptual Sensitivity .....	185
Threshold Setting and False Alarm Rate .....	187
Human Perception of Relative Magnitudes .....	188
5.3 - Display Parameters Affecting Observer Performance .....	190
Luminance .....	190
Resolution .....	191
Regeneration Rate and Viewing Distance .....	191
Operator Training .....	193
Summary and Conclusions .....	194
5.4 - Human Factors Consideration.....	195
Design Philosophy .....	195
Summary of Some Key Human Factors Issues .....	197

5.5 - General Structure of the Image Processor .....	200
5.6 - System Overview .....	207
Fixed Mode .....	208
Free Mode .....	215
Chapter 6 - Radio Astronomy Applications - Molecular Spectroscopy .....	220
Chapter 7 - Geophysical and Material Sciences Applications .....	231
7.1 - Introduction .....	232
7.2 - Acoustic Emission Analysis Techniques .....	236
7.3 - Possible Applications of the Signal Detector in Acoustic Emission .....	239
7.4 - Industrial Applications .....	240
Source Mechanism Studies .....	240
Failure Mechanism Studies .....	241
Leak Detection .....	242
Kaiser and Felicity Studies .....	242
Thermoplastic Studies .....	243
7.5 - Geophysical Applications .....	243
Fracture Density Studies .....	243
Stress Field Analysis .....	245
Source Mechanism Studies .....	246
General Deformation Studies.....	247
Postlude .....	253

## Appendices

A2.1 - The Discrete Fourier Transform of Signals and Noise . . . . .	257
A3.1 - Polarization Synthesis . . . . .	273
A3.2 - Coherency Matrix and Degree of Polarization . . . . .	279
A3.3 - Effects of Noise on the Measured Degree of Polarization . . . . .	280
A3.4 - The Linear Phase Invariance Property of Coherent Power . . . . .	286
A3.5 - Generalized Coherence Mathematical Details . . . . .	287
A3.6 - Generalized Coherence Statistics . . . . .	291
A3.7 - The Possibility of Truly Coherent Detection . . . . .	296
A3.8 - Theory of ANOVA . . . . .	297
A3.9 - ANOVA Performance . . . . .	302
A3.10 - Carrier Wave Detector Weighting Factors and Approximate Detection Statistics . . . . .	320
A3.11 - Cluster Detector Details and Statistics . . . . .	361
A3.12 - An Evaluation of Classical Algorithms in the SETI Context . . . . .	374
A4.1 - Large Memory Systems . . . . .	381
A4.2 - Discussion of Alternative Architectures . . . . .	389
A4.3 - Implementation for (1) Degree of Polarization and (2) Broadband Pulse Detection (with GCV) . . . . .	392
A4.4 - Survey of Image Processors . . . . .	398
A5.1 - Experiments on Observer Detection and Recognition of Simulated Signals in Noise . . . . .	404
A7.1 - Experimental Procedure Outlines for Several Feasibility Studies of Non-Astrophysical Uses of the Multichannel Spectrum Analyzer and Signal Detection System . . . . .	423

## GLOSSARY

A scan refers to one spectrum--8 MHz x 1 Hz, with 16 bits per channel produced in one second.

The entire array refers to 1000 seconds of the MCSA output and is a 1000 sec x 8 MHz matrix of raw data, (1000 scans).

A block is a 20 sec x 40 Hz section of the entire array, 800 points.

The compacted array is a 50 x 250,000 matrix of 9-bit numbers.

A pattern block is a 50 x 200 matrix of 9-bit numbers which is a vertical strip of the compacted array.

A pattern class is a set of 9-bit numbers with specified bits equal to 1 which determine the class and no constraint on the unspecified bits. Thus, both (0, 1, 1, 1, 0, 1, 0, 1, 1) and (1, 1, 0, 1, 1, 1, 0, 1, 0) belong to the class (\*, 1, \*, 1, \*, \*, \*, 1), and also to the class (\*, 1, \*, \*, \*, \*, \*, \*), while (0, 0, 1, 1, 0, 1, 0, 1, 1) belongs to neither of these classes.

A pattern point is any 9-bit number of the compacted array.

A cell is the time-frequency location of a pattern point. Note the one-to-one correspondence between blocks and cells.

In addition, the following mnemonics are used widely throughout the report:

AC - Accumulator

ANOVA - Analysis of Variance

CD - Cluster Detector

CFSR - Common Frame of Semantic Reference

CWD - Carrier Wave Detector

ETI - The Extraterrestrial Intelligence

GCV - Generalized Coherence Value

MCSA - Multi-Channel Spectrum Analyzer

NBIT - Numerical Battery of Independent Tests

OSD - Oasis Signal Detector

PD - Pulse Detector

PROM - Programmable Read-Only Memory

RFI - Radio Frequency Interference

RAM - Random Access Memory

**ROM - Read-Only Memory**

**SETI - Search for Extraterrestrial Intelligence**

**SNR - Signal to Noise Ratio**

**Gbit - Gigabit,  $10^9$  bits**

**Mbit - Megabit,  $10^6$  bits**

**Byte - 8 bits**

**Chapter 1**  
**Introduction**

**The Search for Extraterrestrial Intelligence (SETI)**

You may seek it with thimbles--and seek it with care;  
You may hunt it with forks and hope;  
You may threaten its life with a railway share;  
You may charm it with smiles and soap--

We have sailed many months, we have sailed many weeks,  
(Four weeks to the month you may mark),  
But never as yet ('tis your captain who speaks)  
Have we had the least glimpse of a Snark!

--Lewis Carroll, The Hunting of the Snark

Because of continuing advances and discoveries made by astronomers in the "outer" universe and by microbiologists in the "inner" universe, our estimates of the probability of life being widespread in the universe are growing rapidly. At the same time, technological advances are increasing the probability of our being able to successfully communicate over the large interstellar distances that might exist between adjacent intelligent civilizations. That this idea has clearly emerged from the realm of science fiction into that of science fact may be seen from a statement made by the National Academy of Sciences in their Recommendations for Astronomy in the 1970's.

"Our civilization is within reach of one of the greatest steps in its evolution: knowledge of the existence, nature and activities of independent civilizations in space. At this instant, through this very document, are perhaps passing radio waves bearing the conversations of distant creatures--conversations that we could record if we but pointed a telescope in the right direction and tuned to the proper frequency.

"More and more scientists feel that contact with other civilizations is no longer something beyond our dreams but a natural event in the history of mankind that will perhaps occur in the lifetime of many of us. The promise is now too great, either to turn away from it or to wait much longer before devoting major resources to a search for other intelligent beings.

"In the long run this may be one of science's most important and most profound contributions to mankind and to our civilization."

### 1.1 - The Evidence for Life

Radio Astronomers have found increasingly large organic molecules within gas and dust clouds distributed throughout our galaxy. Molecules such as formaldehyde, ethyl alcohol, formamide, and others with molecular weights up to 123 have been found.

We have artificially created the building blocks of life--amino acids, nucleic acids and primitive proteins--in our laboratories by supplying energy to simple mixtures of water, ammonia, and methane. The energy has been successfully supplied by a great variety of sources, such as electric spark, ultraviolet radiation, heat, acoustic waves, x-rays, and nuclear particle bombardment (Ponnamperuna, 1972). Since these simple substances and energy sources are likely to be present on any newly formed planet, the precursors of life probably are widespread, and although we have no direct evidence that further development or evolution will necessarily occur, it is certainly a possibility. These same building blocks have also been found in moon dust and in meteorites that have landed on the Earth. This indicates that Earth has no monopoly on the organic compounds necessary for life as we know it.

In all our studies of nature, we have yet to encounter a truly unique object or event. Nature tends to create things of many kinds, but each kind is well-populated. We find the same elements in our Sun and in all the other stars and in all the distant galaxies as we do right here on Earth. We



observe that the same laws of physics hold everywhere in the universe. The Sun and the Earth are but average members of our galaxy. There is no basis to believe the Earth is in any sense unique.

## 1.2 - Where Is Extraterrestrial Life to Be Found?

Our space probes to date have indicated little hope for life on the Moon, Mars, or Venus, and conditions on the remaining planets of our solar system seem hostile to any life. Moving out to the stars in our galaxy, we can expect to find life-bearing planets associated with only certain types of parent stars.

The parent star should not be too large, since such stars burn themselves out in too short a time for intelligent life to evolve. Nor should it be too small, since the planet then has to be so close to get enough heat that tidal effects tend to lock the planet with one side always facing the star, similar to Mercury and the Moon. This causes all the atmosphere to freeze out onto the cold side. The parent star should be a slow rotator, indicating that it may have lost most of its initial angular momentum to a planetary system. In our solar system, 98% of the total angular momentum is carried by the planets, and the Sun is a very typical slow rotator.

The heavy elements necessary for life are formed only in the cores of stars. They are not present in the very old first generation stars, but rather are present in the younger second and subsequent generation stars that have formed out of the debris of their exploded forefathers. Many stars are part of small groups of two, three or more stars rotating about a common center of gravity. Stable planetary orbits are less likely to occur in such multiple star systems.

3

All of the stellar characteristics described above can be measured from the Earth, enabling us to choose which specific stars are most likely to harbor life. The factor which takes very unlikely individual occurrences, and makes them likely in an overall sense, is the vast size of our universe. Thus, if we eliminate from further consideration all stars not meeting all of the above criteria, we are still left with around 20 million good candidates in our own galaxy.

### 1.3 - How Should We Look for Extraterrestrial Life?

Considering the time and energy involved, the possibility of actual interstellar travel seems very remote. That leaves us with attempting to converse from afar, probably by some form of electromagnetic radiation. We will have to ignore for the present the fraction of civilizations that might be sending out gravity waves or other signals as yet unknown to us.

The two obvious candidates for electromagnetic signalling are light and radio waves. The advent of the laser adds considerably to the possibility of light signalling, but one basic difficulty is that of trying to outshine the parent star. Stars radiate most strongly in just those spectral regions where light signalling is apt to occur. Another difficulty is that the diffuse material between the stars tends to absorb light, making such signals be rather short-range.

Radio waves, on the other hand, are not bothered appreciably by stellar interference, and are regarded as being the most likely signals to look for, by most workers in the field. Our own present capabilities are quite formidable. If we took our most sensitive receiver, our most powerful transmitter

and our biggest antenna and combined them in an interstellar communications link, we could communicate with a twin of that system at a distance of 100,000 light years. That fact in itself goes a long way toward making the whole idea of interstellar communications seem possible.

Our civilization is an "emerging" one from the standpoint of interstellar communications. We have not yet matured to the point where we consider it practicable to send out signals with no hope of reply for hundreds of years. On the other hand, we could receive whatever signals already exist with much less effort, and with the possibility of immediate success. The value of such one-way communications should not be discounted. Our own radio and television broadcasts, as well as books, are one-way communications in the sense that we cannot immediately reply to the sender. Our knowledge of early Earth civilization through archaeological exploration comes through a communication system that is not only one-way, but is also irreversible and highly unintentional. Yet it is of inestimable value to us.

There are two types of radio signals that we might search for. There are those intended for the internal use of the sender, analogous to our own radio, TV, radar, etc., transmissions which inevitably leak out into space. Then there are those intended specifically to attract attention or serve as a space beacon of some sort. Of the two, leakage signals are expected to be much weaker, since they represent wasted power from the sender's viewpoint. Leakage signals are also more difficult to search for, since we cannot make reasonable assumptions as to expected frequencies, modulation methods, bandwidths, etc. We have only to consider the chaotic appearance

Earth leakage signals would have to an outsider to appreciate the magnitude of this problem (Sullivan, 1978 ). Beacon signals, on the other hand, can be expected to be optimized in the sense of making it easy for us to find them, since that is their purpose. To be sure, we should not forget about leakage signals, but their discovery is more likely to be accidental.

#### 1.4 - Anti-Cryptography vs. Anthropocentrism

Since it is not possible to thoroughly search the entire dimensions of frequency, time, space and signalling methods, it is necessary to construct a search strategy that will reduce or eliminate as many of these dimensions as possible. This involves making a number of assumptions about the characteristics of the signals we are looking for.

Our search strategy is based, in so far as possible, on physical laws or properties which are true throughout the universe. The radio frequency spectrum can thus be limited on the lower end by galactic noise at about 1 GHz, and on the high end by quantum noise at about 100 GHz. Dispersion in the interstellar medium limits useable bandwidth to a few KHz. These limitations should be recognized by any civilizations considering interstellar communications. Further, communication theory suggests very slow signalling rates and narrow bandwidths for maximum range.

The energy-conscious designer of an interstellar beacon will choose all its operating characteristics in such a way as to maximize its probability of detection. This means that if a search dimension can be minimized or eliminated by a design choice, it will be. An example of this

is the dimension of time. If a beacon transmits in different directions at different times, we would be lucky to be looking toward it at just the time it was transmitting toward us. On the other hand, if a beacon transmits continuously and omni-directionally, we need not concern ourselves with the unknown of when to look at it, and the dimension of time is effectively eliminated.

Thus, one approach in the design of a signal detector is to develop an idea of what sort of signal would constitute an optimal interstellar beacon, in terms of distinguishability from natural sources, as well as providing ease of reception. On such grounds, one conclusion (Dixon, 1973) is that the signal is in the 1-3 GHz range, circularly polarized (with binary left and right hand modulation), narrow band, and possibly Doppler corrected. For the reasons above, the Oasis signal detector, as seen in the next two chapters, is especially sensitive to narrowband polarized signals.

In the event that such a signal is not the optimal beacon, or if the transmitted signal is nonoptimal, the OSD is designed to be sensitive to a broad class of signals, including those transmitted routinely and accidentally from Earth into space.

### 1.5 - Why Should We Look?

Consider the continuum of possible development levels of civilizations, and their corresponding communications capabilities. Our civilization has only very recently achieved the technology necessary for interstellar communications. We cannot communicate with less developed civilizations,

because they cannot construct their half of the circuit. We can communicate (with difficulty) to other civilizations of equal development to our own, and can do so with increasing ease to civilizations more and more advanced than ours. If we assume that we are midway along the infinite time line of development, then we cannot reach all those others in the less developed half. The chances of finding another civilization of equal development to our own is small, since that is analogous to finding a single point on an infinitely long line. This leaves only the more advanced civilizations as being those most likely to be found.

We know that our own civilization is capable of destroying itself through wars, overpopulation, pollution and other causes. It is possible that all civilizations suffer such a fate at some critical point in their development. Perhaps that is a general law of nature; we simply do not know. If this were the case, we might never find another civilization. On the other hand, if we were to find another civilization (that would very likely be more advanced than we), this would be evidence that it is at least possible for a civilization to attain maturity without destroying itself. This might provide the impetus we need to redouble our efforts to eliminate such possibilities. Perhaps the signals themselves would tell us how to solve these problems. This might be referred to as a cosmic feedback effect, where the transfer of knowledge from one civilization to another tends to stabilize, preserve and synchronize them.

Our sense of perspective would change if we realized we were not alone in the universe. Perhaps some of our local problems would not seem quite so important. Looking at things from a larger viewpoint might reduce the foolish, jealous and chauvinistic tendencies we have.

Mankind needs a continuing challenge. Without it, cultural stagnation will set in and our civilization will decline. There must always be new frontiers and adventures to face or we will sit back on our laurels and let the universe go by.

#### 1.6 - How Can We Understand Their Messages?

Whatever they say, it will be said in such a way as to make it as easy as possible for us to understand it. Linguists say that a Common Frame of Semantic Reference (CFSR) must exist before meaningful communications can occur in any context. In other words, mutually common knowledge or experiences must serve as the basis for initial understanding. Once such a basis is established, knowledge which is known only to one can be imparted to the other. We have successfully translated ancient lost languages, but a large CFSR exists between any two races of humans. Dolphins are apparently quite intelligent, and have complex speech, yet we have been unable to communicate with them. The CFSR is less obvious in that case.

An obvious CFSR that we must share with any intelligent technological civilization is that of counting and numbers. This has been exploited by Hans Freudenthal (1960), a Dutch mathematician, who created Lingua Cosmica. The basic idea is to send a repetitive message which begins with a counting sequence and builds up logically by telling how to answer a transmission or telling where other transmissions might be found. The first message might be referred to as an entrance examination.

. .. .... (etc.)

Each subsequent message or lesson introduces exactly one new idea or symbol, using it in many different contexts with previously defined symbols, to make its meaning obvious. For example, the concept of "greater than" can be introduced by using a new symbol ">" in context with the numbers learned in the entrance examination, i.e.,

2?1 3?1 3?2 4?1 4?2 4?3 (etc.)

Other schemes have also been created (see Arbib, 1979).

The important thing to note here is not that we really expect interstellar signals to be exactly like Lingua Cosmica, but that at least methods seem to exist for information to be sent, received, and understood among different civilizations.

### 1.7 - What Is Being Done Now?

To date, we have only scratched the surface of what we are capable of doing, in terms of actually looking. Contrary to popular belief, radio observatories do not spend all their time searching for intelligent signals. It would be closer to the truth to say that radio observatories never search for intelligent signals. The actual situation is that most observatories never search at all, a few search for a tiny fraction of their time, and only one searches all the time. A long term search for pulsed signals is being conducted in the USSR, using very small, non-directional antennas. A full-time search for narrow-band signals using a large radio astronomy antenna has been in operation since 1973 at Ohio State University. The various searches to date are summarized in Table 1.1 (excerpted for data provided by Jill Tarter of NASA-Ames Research Center).



Table 1.1 - Past and Present Search Projects

<u>Date</u>	<u>Observers</u>	<u>Location</u>	<u>Telescope size (meters)</u>	<u>Frequency range (MHz)</u>	<u>Objects Searched</u>	<u>Total time</u>
1960	Drake	West Virginia	26	1420	2 stars	2 weeks
1968-69	Troitskii, Gershtein, Starodubstev, Rakhlin	USSR	15	927 and 1420	12 stars	1/2 day
1970 to Present	Troitskii, Bondar, Starodubstev	USSR	1/4	600, 927 and 1863	All-sky, pulses	4-1/2 years
1971-72	Verschuur	West Virginia	91	1420	9 stars	1/2 day
1972-76	Palmer, Zuckerman	West Virginia	91	1420	600 stars	3 weeks
1972 to Present	Kardashev	USSR	1/10	1337-1863	All-sky, pulses	unknown
1973 to Present	Dixon, Ehman, Arnold, Kraus, Raub	Ohio	53	1420	All-sky	6 years
1974 to Present	Bridle, Feldman	Canada	46	22,000	500 stars	1 week
1975-76	Drake, Sagen	Puerto Rico	305	1420, 1653, 4 galaxies and 2380	4 days	4 days
1975-79	Isreal, De Ruiter	Netherlands	100	1415	250 stars	2 weeks

1976 to Present	Bowyer, Welch, Tarter, Zeitlin	California	26	1420 and 1670	All-sky	3 months
1976	Clark, Black, Cuzzi, Tarter	West Virginia	43	8500	4 stars	1/3 day
1977	Black, Clark, Cuzzi, Tarter	West Virginia	91	1666	200 stars	4 days
1977	Drake, Stull	Puerto Rico	305	1666	6 stars	1/2 day
1977 to Present	Wielebinski, Seiradakis	Germany	100	1666	6 stars, pulses	2 hours
1978	Cohen Malkan	Puerto Rico, Massachusetts, Australia	305 36 63	1666 22,000 1612	25 star clusters	3 days

All of these searches basically employ previously-existing radio-astronomical equipment and techniques. No positively identified signals from extraterrestrial civilizations have been found, although several single-event unexplained cases have been noted (Kraus, 1979). Numerous papers and books have been published on the subject, and many national and international conferences have been held. A scientific journal devoted exclusively to the subject (called Cosmic Search, Kraus et al., 1979) is now being published.

To give proper perspective to what is being done now in comparison to what could be done, consider the search for earthquakes in the state of California. An easy earthquake search experiment that can be performed without specialized equipment is to simply stand outside somewhere in California for 10 minutes and note carefully if you feel any vibrations. The most likely outcome of this experiment is not indicative of the true situation, simply because its scope is too small.

In 1971, NASA sponsored Project Cyclops (Oliver, 1971), a feasibility study of what would be required in terms of time, money, and personnel to mount a full-scale search for intelligent extraterrestrial radio signals. The basic recommendations were to construct several 100 meter diameter dish antennas, and search the nearby stars in great detail. If no signals were found, additional dishes could be constructed in stages, up to a limit of about 1000, to search progressively more distant stars. If a signal were found at an early stage, the entire array would not need to be constructed.

In 1975-76, NASA held a series of Science Workshops on Interstellar Communications (Morrison, et al., 1977). A wide variety of well-known

scientists in various fields were convened to re-examine the SETI concept in light of current knowledge and technology. The consensus of all those attending was that it was both timely and feasible to begin a serious search for extraterrestrial intelligence and that it can be done now at only modest cost.

#### 1.8 - Detection Techniques Used in Previous SETI Projects

All current and previous SETI projects use equipment that is either simple and inexpensive, or that has been adapted from some other original purpose. The techniques may be broadly grouped into two categories-- those which search for wideband pulsed signals and those which search for narrowband continuous signals. Refer to Table 1.1 for more details on these various searches.

##### Wideband Pulsed Signal Techniques

Two projects in the USSR employ widely separated (thousands of miles) receivers, with simple non-directional antennas, to make tape recordings of pulsed signals. The tapes are later sent to a central location and cross-correlated to determine if any time-coincident pulses were observed among the various receivers. Typically each receiver observes many random pulses due to RFI, lightning, and other phenomena local to each receiver. Any such pulses would not be observed at the other receivers however, so they are eliminated by the cross-correlation process. Pulses which are received simultaneously at all or most of the receivers may be inferred to have a common, most likely extraterrestrial, origin.

A second technique for finding pulses is the use of pulsar search equipment, being done in Germany. Pulsars are natural astronomical objects

that emit periodic broadband pulses, which are received as pulses that sweep downward in frequency with time. The apparent frequency sweep is caused by the velocity of wave propagation in the interstellar medium being a function of frequency (i.e., high frequency signals travel faster than low frequency signals). Pulsar detection equipment typically uses several receivers spaced in frequency, to observe both the pulsing and sweeping phenomena.

### Narrowband Continuous Signal Techniques

The great majority of projects search for narrowband signals, from specific target objects. Their techniques differ from one another principally in the method and degree of achieving narrow bandwidth, and in the number of receiver channels in simultaneous operation. In virtually all cases, the equipment used is adapted from radioastronomical natural line detection apparatus.

Banks of analog filters are used to achieve moderate bandwidth (10 kHz) and a moderate number of channels (50). To achieve a larger number of channels (1000) autocorrelation receivers are used. These receivers digitally compute the Fourier Transform of the autocorrelation function of the incoming signal in real time, giving as output the power spectrum of the signal. For extremely narrow bandwidths (10 Hz) and high number of channels, the incoming signal is recorded on a high-speed tape recorder for later (non-realtime) processing. The tapes have been analyzed by a variety of techniques, including direct Fourier Transformation in a large mainframe computer, optical Fourier Transformation using a laser and lens system, and Fast Fourier Transformation using a mini-computer. In all cases, the power spectrum is computed and inspected for peaks.

The project in Ohio employs a number of specialized SETI signal detection algorithms. The sky is continuously swept, at the Earth's rotation rate. In that mode of observation, the intensity of signals which come from small-angular-extent objects appears to trace out the antenna pattern, as a function of time. This fact is used by continuously computing the normalized cross-correlation function of the received data with the antenna pattern, in each channel. This procedure discriminates against large diameter natural astronomical objects and short-time peaks. In addition, adjacent channels are intercompared and if a detection occurs simultaneously on two or more channels, it is ignored. This discriminates against wide-band signals. When a signal is detected in one of the channels, a narrowband tuneable filter is automatically activated and swept slowly across the frequency range of the detection channel, to provide a re-examination of the signal at 1/10 the discovery bandwidth. RFI is largely eliminated through use of a beam-switching technique, where the receiver output is the difference between two closely-spaced beams on the sky. Any real source appears in only one of the beams at a time, and hence is observed normally. RFI in general is received in the sidelobes of the antenna, which are the same for both main beams. Thus RFI is effectively subtracted out.

Several other programs make use of data and observations done for other radioastronomical purposes, which are then reanalyzed with SETI in mind. This is the case for the projects in the Netherlands (examination of radio sky maps at known star positions), in California (a parasitic receiver that makes a separate recording of narrowband signals, from whatever the radio telescope happens to be looking at), and the recent one by Cohen and Malkam (examination of data originally taken to find natural lasers).

## 1.9 - The Next Step

The next quantum jump in signal sensitivity requires the use of receiving equipment designed and constructed specifically for SETI. This is analogous to building seismographs to detect earthquakes in California. Existing radio telescope antennas are quite adequate for the present; only their receivers need be changed. Fortunately, receivers are much cheaper than antennas.

Before such specialized receivers can be constructed, they must be designed. Even before that, they must be conceptualized. Decisions must be made as to what types of signals to expect, and optimum search algorithms designed to detect them. Such signals might be from beacons. At the same time, means must be devised to catch those signals which are unpredictable, amorphous or strange. Such signals might be leakage. Careful searching takes a long time, requiring full-time operation at several observations around the earth, for many years. This is analogous to operating the seismograph for ten years, instead of ten minutes.

Overlaying all extraterrestrial signals are the overwhelmingly powerful signals being generated by our own radio and television stations. The SETI receiver must operate in the presence of such interference constantly. Even worse than the signals generated here on Earth are those transmitted by artificial satellites, for they look directly down the throat of the SETI antennas. The number of transmitters in operation, both on Earth and in space, is growing rapidly with time. Eventually this will make it impossible to conduct SETI programs anywhere on the surface of the Earth. The alternative is to construct SETI antennas in space or on the moon. This will be much more expensive and could not occur for many decades. Thus there is a great urgency to proceed now.

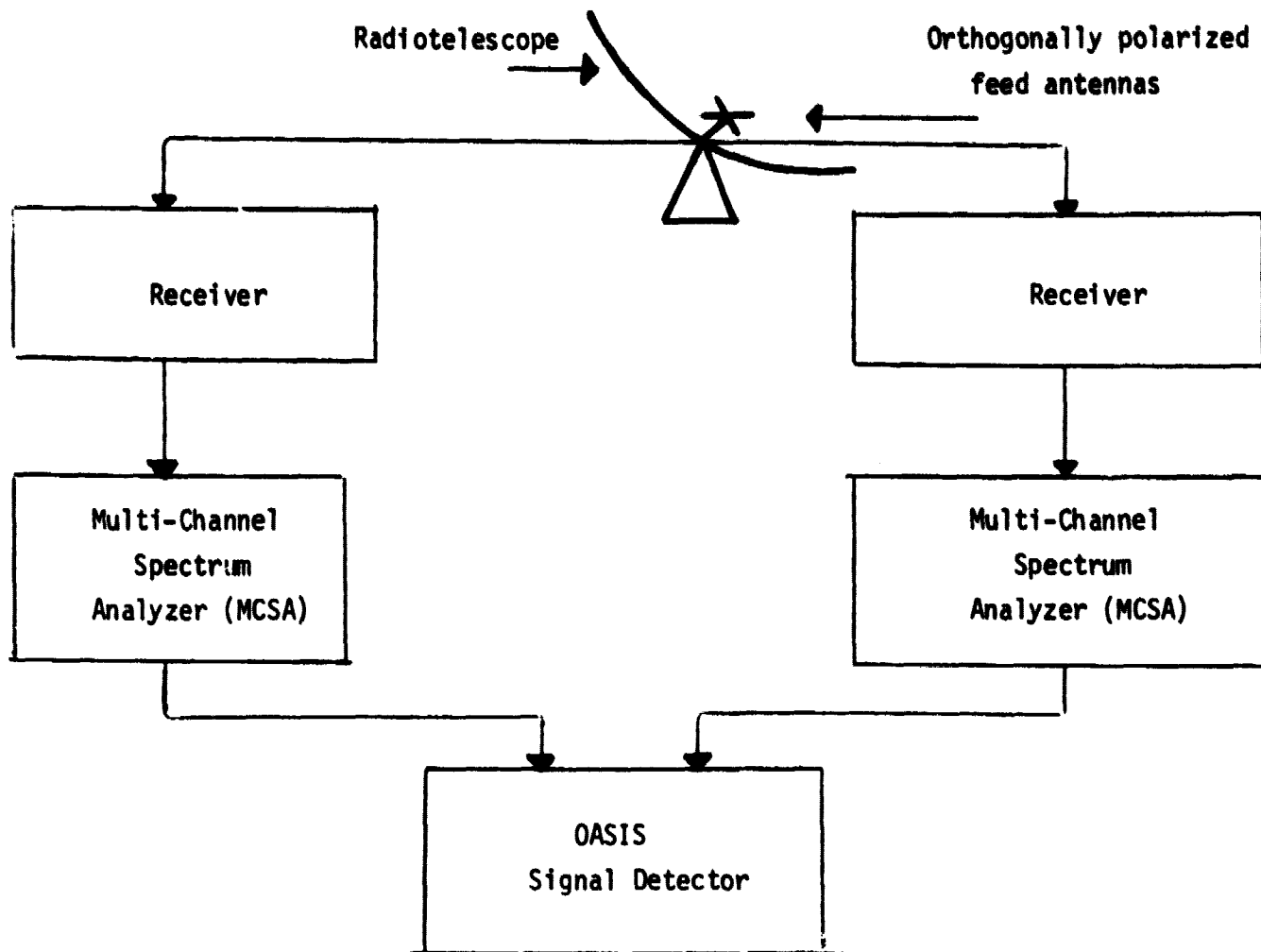
## 1.10 - The Receiving System

Figure 1.1 shows conceptually the components which make up the SETI receiving system. A radiotelescope containing two orthogonally polarized feed antennas is connected to two separate receivers. The output of each receiver is fed to two separate Multi-Channel Spectrum Analyzers (MCSA's). The MCSA is a special purpose SETI device, capable of analyzing 8 MHz of bandwidth with a resolution of 1 Hz. The 8 MHz input bandwidth is digitized and then initially split into 120 bands using a digital bandpass filter bank. These bands are then subdivided to a resolution of 1 Hz by 120 microprocessors that continuously compute the Discrete Fourier Transform of the filter outputs. Both the input and output data is in complex form (i.e., possessing both magnitude and phase).

Each word is 16 bits in length, but since we use only 4 as being of sufficient accuracy, then the output data rate is 16 megabytes per second. This data is fed to the Oasis Signal Detector, where it is searched for possible intelligence. In actual operation, the radio telescope examines a target star for up to 1000 seconds. All of the data for a given target is to be available for computation at one time, and a total of 16 Gigabytes of data is stored digitally.

To give perspective to the magnitude of this task, consider the following analogy. 16 megabytes per second is about one encyclopedia set per second. These encyclopedias do not contain words and sentences, but are filled with random letters and numbers. As we read through these encyclopedias (after having taken a speed reading course), we have no idea when an intelligent message might begin (today or 5 years from now), or in what language it will be. To make matters worse, each encyclopedia will





System to Seek Extraterrestrial Radio Signals

Figure 1.1

only contain one word of the message (average of 8 bytes). For example, the first set of encyclopedias might contain "The", and the second set "message", and the third "is", and the fourth "zk3ja". As usual, just when you get to the good part, the data got garbled by noise, but this serves to illustrate the real fact that noise will always make the signal less clear than it otherwise might be.

To make life simpler, we assume that the message will always occur in a consistent volume and on a consistent page, once it gets started, for example, it might always be in the J volume on page 258. Of course we don't know which volume or page it's on, so we still have to search everything, but this assumption allows us to reorganize our search. We use a special device called an MCSA that rips all the pages out of each encyclopedia as it appears, sorts them out by volume and page number, and rebinds them again (all in one second). Now we have a new set of encyclopedias to read, but all the page 258's of all the J volumes now are bound together in sequential order. We still have just as much reading to do, but now we don't have to remember as much.

Occasionally the random letters and numbers will appear to say something purely by accident. The analogy here is that enough monkeys at enough typewriters must inevitably produce strangely real phrases. We call these false alarms. They are expensive because they force us to stop everything and investigate their source. We try to avoid them if possible, but we have to be careful that in doing so we don't miss the real signal.

We are also bothered by occasional pages from the Readers Digest that are pasted right over some of our encyclopedia pages. They correspond to our receiver picking up terrestrial radio and television stations. These

pages are usually easy to spot and skip over, but they are distracting and they cover up an encyclopedia page that might have contained the message.

The last complication is the fact the data in reality comes from two antennas of different polarizations. This is analogous to splitting the encyclopedia in two, with all the letters in one half and all the numbers in the other half. We don't know if the message will be all letters, or all numbers, or some combination of the two. We have to read both the letters and numbers simultaneously, trying to intersperse them appropriately to make up the entire message.

#### 1.11 - Purpose of This Study

This study is a design of that part of the specialized SETI receiver discussed above that actually finds and recognizes the signals. The ideas and designs incorporated in this system will undoubtedly be refined and modified by future workers. Eventually a system will actually be constructed. This is the beginning of the next step.

#### 1.12 - The Oasis Signal Detection System

Three separate signal detection philosophies are embodied in the Oasis system, illustrated conceptually in Figure 1.2. The Carrier Wave Signal Detector seeks the extreme case of a signal that has nearly zero bandwidth, and may be drifting slowly with time. The Pulse Signal Detector seeks the other extreme case of a signal that has a broad bandwidth, and may be pulsed in time. The Battery of Tests seeks all those signals in between the extreme cases; those that are unpredictable.

Each of these detection methods will be explained in detail in the chapters to follow. Chapter 2 discusses the philosophical differences between

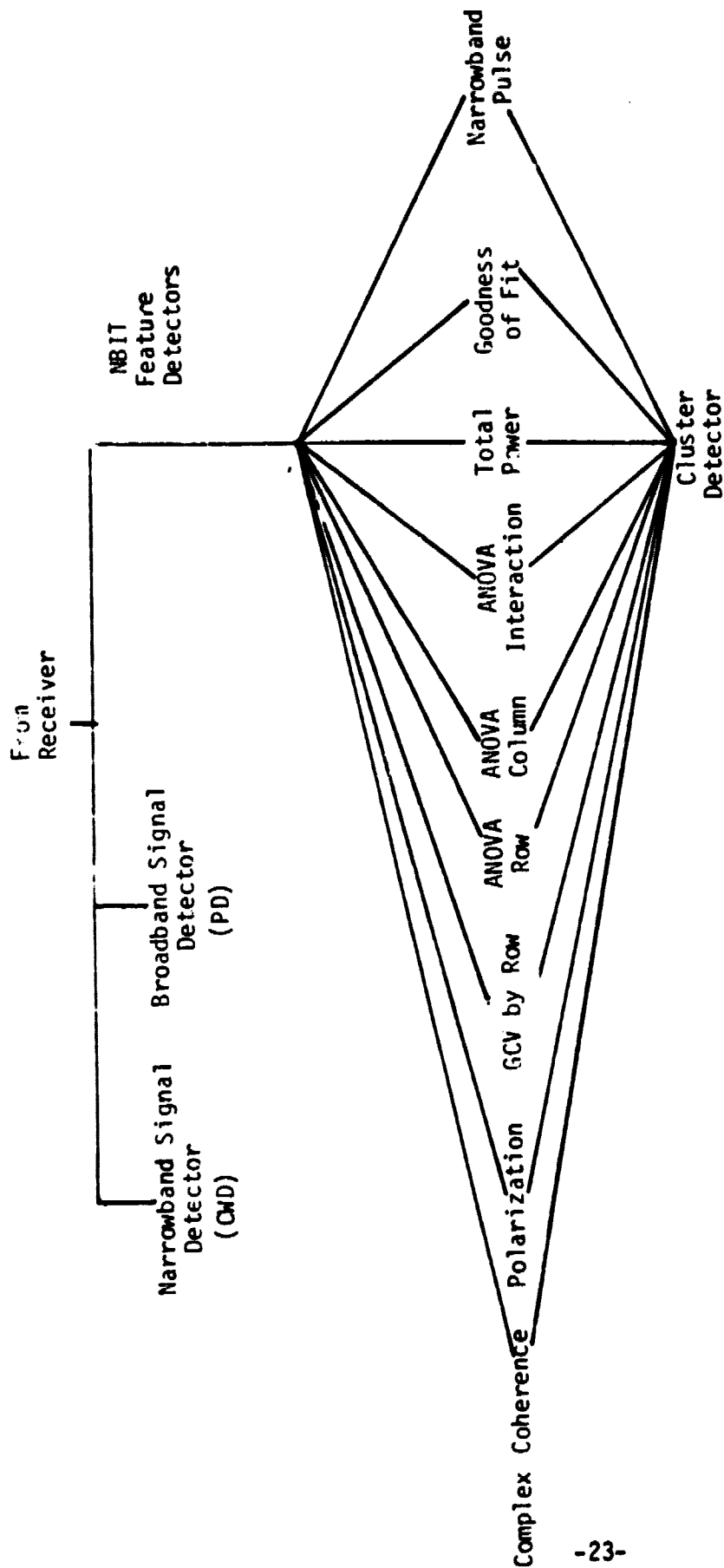


Figure 1.2 - Oasis System Signal Detectors

signals and noise. Chapter 3 lays the detailed mathematical bases and philosophies of the Oasis detection algorithms. Chapter 4 summarizes the overall Oasis system in block form, and provides hardware implementation details. Chapter 5 focuses on the unique capabilities of the human being to detect unexpected signals, as well as the role of the operator in the Oasis system. Chapter 6 discusses the radio astronomy applications of the system. Chapter 7 provides illustrations of non-astrophysical areas where the Oasis system could be put to use in solving a variety of current scientific problems.

## References

1. Arbib, M.A.; Minds and Millernia: The psychology of Interstellar Communication; Cosmic Search, Vol. 1, No. 3, Summer 1979, p.21.
2. Cameron, A.G.W., Interstellar Communication--The Search for Extraterrestrial Life, W.A. Benjamin, New York (1963).
3. Dixon, R.S., A Search Strategy for Finding Extraterrestrial Radio Beacons; Icarus, 20, 187 (1973).
4. Freudenthal, H., Lincos, Design of A Language for Cosmic Intercourse, North-Holland Pub. Co., Amsterdam (1960).
5. Kaplan, S.A., Extraterrestrial Civilizations--Problems of Interstellar Civilizations, NASA TT F-631 (1971).
6. Kraus, J.D., Gearhart, M.R., and Dixon, R.S., Cosmic Search Magazine, published by Cosmic-Quest, Inc., P.O. Box 293, Delaware, Ohio (began in Jan. 1979).
7. Kraus, J.D., We Wait and Wonder, Cosmic Search, Vol.1, No.3, Summer 1979, p.31.
8. Morrison, P., Billingham, J. and Wolfe, J., The Search for Extraterrestrial Intelligence, NASA publication SP-419 (1977).
9. Narasimha, M., Peterson, A., and Narayan, S., Implementation of Real-Time Digital Signal Processing Systems, SPIE, vol. 154 Real-Time Signal Processing, p. 81 (1978).
10. Oliver, B.M., Project Cyclops--A Design Study of a System for Detecting Extraterrestrial Intelligent Life, NASA publication CR 114445 (1971).
11. Ponnamperuna, C. and Cameron, A.G.W., Interstellar Communication: Scientific Perspective, Houghton-Mifflin, Co., Boston (1974).
12. Ponnamperuna, C., The Origins of Life, Thames and Hudson, London (1972).
13. Sullivan, W.T., Brown, S., Whetherill, C., Eavesdropping: The Radio Signature of the Earth, Science (January 27, 1978), pp.377-387.

## **Chapter 2**

### **Signals and Noise**

Everything gets written down, formally, so that you know at all times where you are, where you've been, where you're going and where you want to get. In scientific work and electronics technology this is necessary because otherwise the problems get so complex you get lost in them and confused and forget what you know and what you don't know and have to give up.

--Robert M. Pirsig, Zen and the Art of Motorcycle Maintenance

## 2.1 - Introduction

The primary purpose of the signal detector is to differentiate between signals and noise. This separation process is based, of course, on the differing characteristics of signals and noise. In this chapter we explore these characteristics, and discuss how they can be used to distinguish one from the other. The actual implementation of specific algorithms to accomplish this separation is discussed in Chapter 3.

## 2.2 - Classifying Signals and Noise

In this section an arbitrary classification of inputs to the signal detector is presented. The classification is based on the origin of the input. The input will be called either signal or noise and the signal can have one of three forms as shown in Table 2.1. It is the purpose of the signal detector to pick out any of the three signals, saving it for subsequent identification, and reject noise.

<u>Classification</u>	
Signals	
	1. Natural - Astronomical
	2. Radio Frequency Interference (RFI)
	3. Unidentified (Includes ETI)
Noise	

Table 2.1 - Classifying Signals and Noise  
of Astronomical Interest



The first signal form comes from natural sources. These include, for example, pulsars, spectral lines from interstellar clouds, the galactic 21 cm line, radio continuum sources, etc. Such natural signals are to be identified and saved as of interest to astronomers.

The second signal form is radio frequency interference (RFI) which is any man-made electromagnetic radiation originating from sources such as radio systems (both ground-based and orbiting satellites), radar, ignition systems, household appliances, etc., depending on the observation frequency. These signals are not usually of direct interest by themselves. However, they are in general sufficiently different from noise that they will be picked out by the signal detector. In many situations it will be desirable to catalog such signals so that later inputs of the same form can be recognized as RFI by the signal identification stage.

Unidentified signals are simply inputs which appear to be different from noise but which cannot be confidently identified as either astronomical or RFI. They may in fact be one of the two, or they may originate from an intelligent source beyond the solar system.

Noise is defined in Section 2.4 below.

### 2.3 - A Global View

In this section we consider the environment in which signals are generated and noise introduced, and how a system to distinguish between the two is structured. Figure 2.1 shows a conceptual block diagram of a complete communications system.

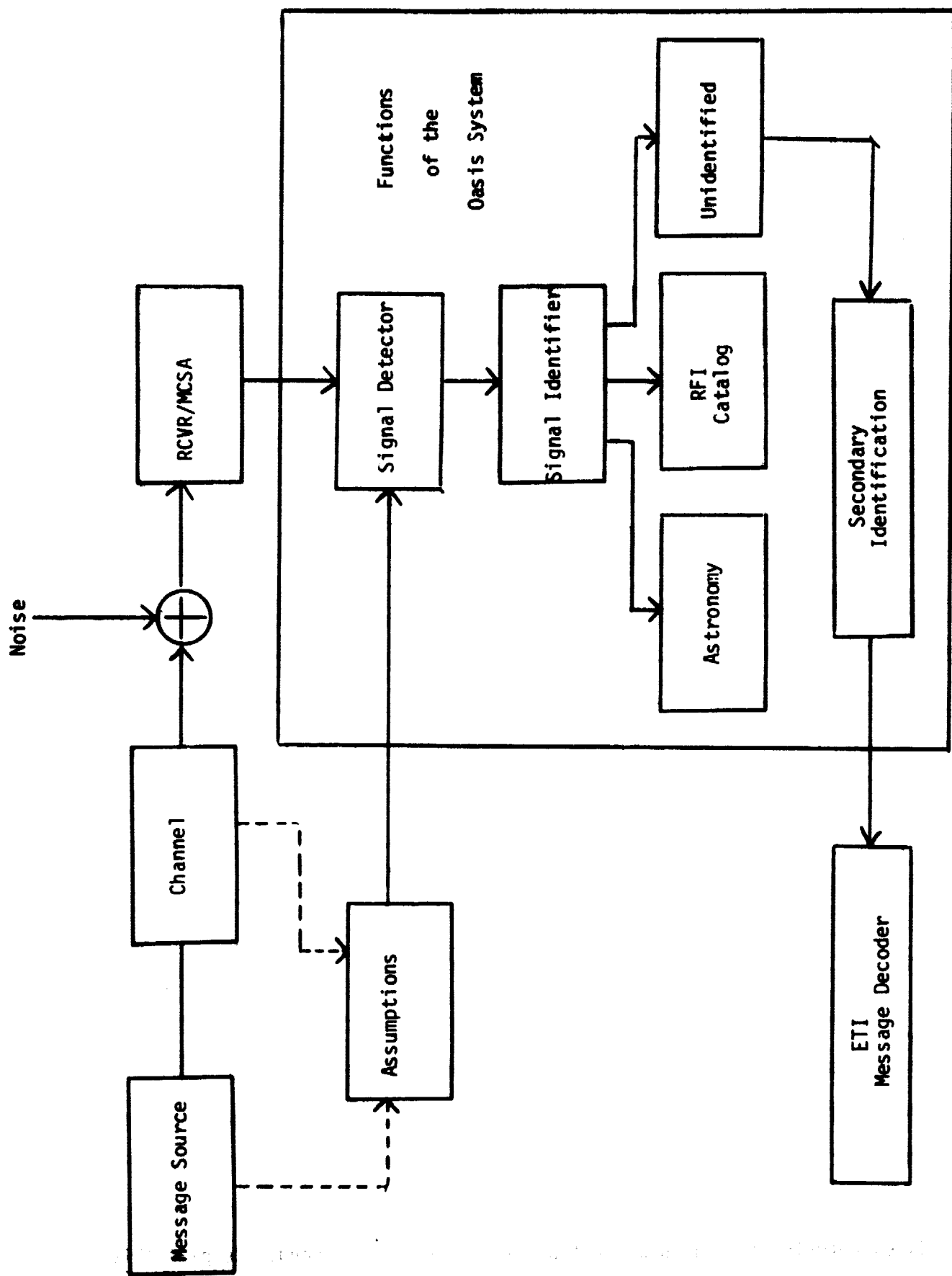


Figure 2.1 - Generalized Communication System

The message source block may represent an ignition system in a nearby automobile, a pulsar or hydrogen cloud somewhere in the Milky Way, or perhaps an intelligent civilization many light years from earth.

The channel represents the medium through which the signal is transmitted. It may be a few miles of atmosphere for a nearby RFI, or scores of light years with components such as atmospheres and ionospheres, interstellar clouds, and nearly empty space. Some of these transmission media can have important effects on the signal as it travels between message source and receiver.

All the noise is assumed to be added to the signal at the receiver. In fact the noise may have come from far out in space, or may have originated in the receiver itself. It is assumed in this study that the receiver and the MCSA already exist. They are not of direct interest to our design except for some control functions associated with the best operation of the signal detector.

The signal detector is the focus of this design project. It has the task of detecting any signals which exist in the expanse of noise which permeates the electromagnetic environment. Its design is dictated by the assumptions which are made about noise and about the various signals. In a classic communication system it is usually assumed that the general character and the statistics of signals and noise are known, and are used to determine the nature of the signal detector. In our case nothing is known about any potential ETI signals, and there is uncertainty about the exact nature of the other signals and the noise which will be received. For this reason the lines entering the "assumptions" block are shown as dashed. Despite this

uncertainty, assumptions must be made because they are an essential prerequisite to the design of the signal detector. These assumptions are presented in Sections 2.4 and 2.5.

The output of the signal detector goes to a signal identifier which has the task of determining whether the signal is natural, RFI, or unidentified. This block is not a primary focus of the present study although it does receive some attention. The signal identifier directs the signal to the appropriate user.

Unidentified signals are stored temporarily or longer for attempted identification by the operator, with the aid of a separate RFI detection and identification systems and various other tools available to the operator.

Finally, of course, we shall want to decode or interpret any ETI signals which are obtained. This problem has been addressed briefly elsewhere (1), and is beyond the scope of this study. If an ETI signal was ever found, resources far beyond those available for the search itself would be brought to bear on its interpretation.

## 2.4 - What Is Noise?

The term "noise" is not uniquely or unambiguously defined. Sometimes it refers to any undesired input. In our case such a definition could be ambiguous. For example, RFI is not of interest or desired for SETI purposes and yet we must initially detect it as a potential ETI signal, since it has many of the same characteristics. Once identified as RFI, we catalog such a signal so as to reject recurrent similar RFI by means of comparison tests. Likewise, astronomical signals are not desired for SETI purposes, but are sent to the astronomical observer for other uses.

For our purposes we define noise in terms of its statistical characteristics. We then define a signal as any input whose characteristics differ significantly from those of noise.

There are many sources of noise, with many different statistical properties, in the electromagnetic environment. However, in the part of the spectrum of interest to us here, the noise is primarily thermal noise generated by the receiver and the Earth's atmosphere. We assume that the noise in the system is Gaussian in distribution, with zero mean and variance  $\sigma^2$ . We further assume that the noise power is uniformly distributed over the spectral range of interest, with a power spectral density equal to  $kT$ , where  $k$  is Boltzman's constant and  $T$  is the equivalent system noise temperature. Hence the noise power is  $kTB$  ( $= \sigma^2$ ) where  $B$  is the bandwidth of the portion of the spectrum under study. Finally we assume that the noise is unpolarized.

## 2.5 - Signal Characteristics

The distinguishing characteristics which are used to identify an input as signal are its:

- a. amplitude
- b. periodicity
- c. polarization
- d. bandwidth

The amplitude of a signal might be used to distinguish it from noise in two ways. First, the average power or amplitude of the signal might greatly exceed that of the noise (a high signal to noise ratio). Then a threshold

detector can indicate a signal whenever some arbitrary threshold is exceeded. An alternative signaling scheme is one in which the average signal power is not necessarily greater than the noise power but the signal power is concentrated into a short time pulse. Such a signal could also be detected by an amplitude threshold device. Of course, Gaussian noise, since it has a theoretically infinite amplitude range, can also trigger such an amplitude detector, although with low probability. The detection process is inherently probabilistic. A detector output caused by noise is a false detection of a "false alarm."

One of the most important characteristics of signals is periodicity, that is their tendency to repeat their form over constant intervals of time. Actually, this periodicity can manifest itself in a number of ways. Many signals have a so-called "carrier" component which is a sinusoidal signal with a well-defined period. Another possible periodicity is a sequence of repeated pulses occurring at some fixed spacing. While it is possible to design a communication system with no periodicity whatsoever, such signals begin to resemble noise and would be very difficult to detect without considerable advance knowledge of their structure. One of the reasons that periodicity is of such great interest to us is that few natural signals exhibit a high degree of periodicity. (Pulsars are a counter-example.)

Signals can never be exactly periodic though they may be so close to periodic that we have no interest in the deviations. We are also interested in signals which have relatively small deviations from periodicity as is true of slowly drifting sinusoids or carriers with very little modulation. Such signals are often called narrowband in that they occupy not a single

line in the spectrum, but rather a fairly small range of the spectrum.

There are a number of ways of detecting periodicity in an input. One is by means of a filter, which might be mechanical, or electrical, or of some other form. If the input is a sinusoid or has sinusoidal components with a frequency near the resonant frequency of the filter, this input or components of the input are passed through the filter. All other signals and noise with frequency components not close to this resonant frequency are "filtered out."

A mathematical approach to isolating or detecting sinusoids is the Fourier Transform, which has the form:

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt \quad (2-1)$$

Suppose that the input  $x(t)$  is a simple periodic signal of the form:

$$x(t) = Ae^{j2\pi f_1 t} \quad (2-2)$$

If  $f = f_1$  in (2-1) the exponents cancel so that the only integrand is the constant  $A$ , leading to an unbounded integral. If  $f \neq f_1$  the exponential function rotates, such that its integral is essentially zero. Hence the effect of the Fourier Transform is to pick out the signal with frequency  $f_1$  and reject all other signals and noise. In our system the signal filtering process is carried out by the MCSA which is programmed to execute a Discrete Fourier Transform (DFT) of the form:

$$X_n = \frac{1}{N} \sum_{k=0}^{N-1} x_k e^{-j2\pi nk/N}, \quad n = 0, 1, 2, \dots, N-1$$

where the  $x_k$  are  $N$  samples of the time function  $x(t)$  over a total record (sampling) period of  $T$  seconds.

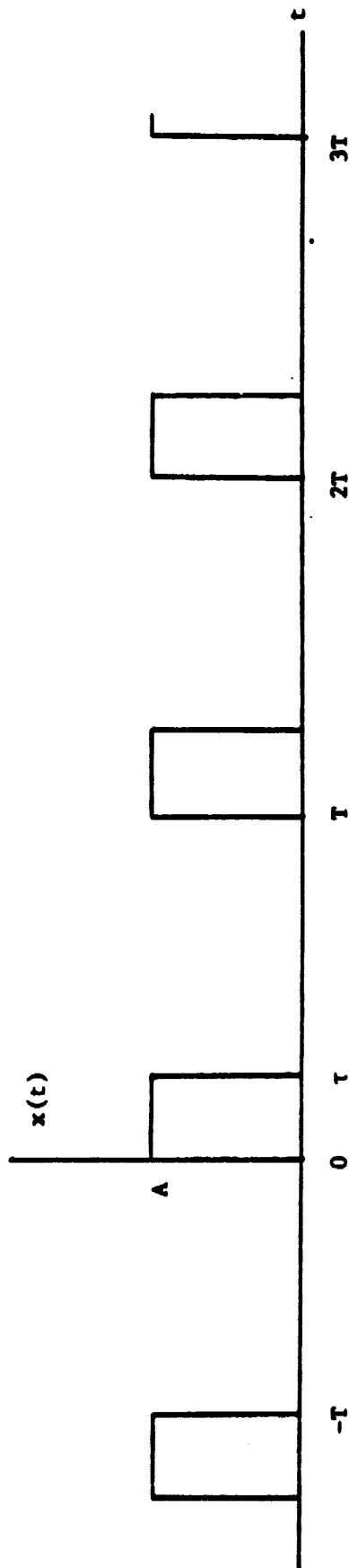
The response of the DFT to both signals and noise is discussed in Appendix A2.1 in some detail. As Equation (A2.1-20) and Figure A2.1.3 show, the response of the DFT to a sinusoid across any one channel has a finite width, following a sinc function law. Hence the MCSA acts like a set of very narrowband filters, which reject signals or noise components outside of a bandwidth of about one Hertz in our case.

Sinusoidal periodicities can also be detected by phase-lock loops which act as a filter with a variable resonant frequency. The loop attempts to adjust its center frequency and remain locked to this frequency as long as the input frequency remains fixed. The loop can follow drifts or shifts in the input frequency, and also angular modulation of the input, as long as the drift rate or modulation rate does not exceed the capacity of the loop.

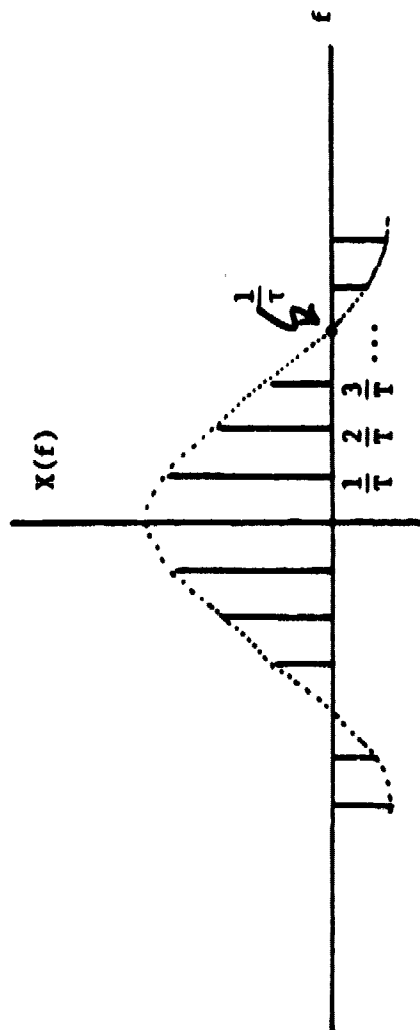
The periodicity may also appear in the form of repeated pulses, such as in Figure 2.2a. The Fourier series coefficients for such a signal vary as shown in Figure 2.2b. The important point here is that there are strong frequency components out to about  $1/\tau$  Hz, where  $\tau$  is the pulse width. That is, the narrower the pulse, the more frequency content there is in the signal. If, for example, the pulse width were about 10 msec, the power in the signal would be spread over about 100 Hz, that is over 100 channels of the MCSA operating with 1 Hz channels. Hence such pulses can be sought by observing many channels at once. In Figure 2.3 we show how a drifting carrier, a pulsed drifting carrier, and a sinusoidal drifting carrier might appear in the noisy output of the MCSA.

The final signal characteristic which we consider is the polarization. We assume here that ETI signals are strongly polarized and noise unpolarized. Hence the presence of unusually high power in an input which lacks polarization tends to suggest an unusually high noise peak. On the other hand, some





a. Time Plane



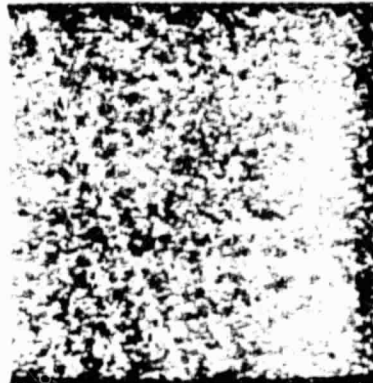
b. Frequency Plane

Figure 2.2. Repeated Pulse Signal

ORIGINAL PAGE IS  
OF POOR QUALITY

Figure 2.3 — Simulated signals in noise. Each display below represents a set of vertically stacked frequency spectra. Each spectrum has been taken at sequential moments in time. Brightness indicates detected power at a frequency. The displays are dominated by noise, while the linear features indicate the detection of carrier waves.

- (a) Carrier wave drifting in frequency  
at a constant rate in time



- (b) Pulsed drifting carrier wave



- (c) Sinusoidally drifting carrier  
wave



polarization in an input which does not have high power tends to suggest the presence of a relatively weak signal in noise.

In this section we have considered three major characteristics of signals: amplitude, frequency, and polarization. Next we use these characteristics to define what kind of signals we will attempt to detect.

## 2.6 - Signal Type Assumptions

The signal detector is designed to detect five different general classes of signals.

1. Monochromatic (Sinusoidal)
2. Drifting Sinusoid
3. Repeated Pulses
4. Narrowband Signals
5. Patterned Signals

These classes are not all mutually exclusive, but the classification is valuable as a guide in setting up signal tests or algorithms. Some tests may detect more than one class of signal.

Monochromatic or single-frequency sinusoidal signals are a very strong possibility in the SETI environment. Such signals could be used as a beacon to attract attention. Or they might be carriers from a radio signal, observed in an eavesdropping mode (2). Finally, they might be very slowly drifting signals with a frequency which appears to be constant over the observation time. In any event such signals would be of great interest and would require attention.

A drifting sinusoid could easily arise from any of the mechanisms in the above paragraph with the added effect of a change in the doppler shift due to planetary motion (ours and theirs). Because such an effect is

expected, and because the sinusoid is relatively simple to transmit and detect, this signal is given perhaps the highest a priori probability of existence, and hence a substantial effort is made to detect it.

Repeated pulses, modulating a carrier in the spectral search range, have the form shown in Figure 2.2a. They could arise for at least two reasons. First, an ETI might reason that instead of transmitting a sinusoid at some given power level it might be better to transmit a higher burst or pulse of power for a short time, and then turn the transmitter off between pulse repetitions. A peak power detector would have an easier time observing such a signal than a sinusoid. A second way in which this signal might arise is from a sinusoidal-signal beacon which is swept across the sky in an attempt to attract attention over a wider range than a focused beam. Such a swept beacon would appear to any one observer as a pulsed signal.

Narrowband signals have their power concentrated over a relatively narrow range of the frequency spectrum. They could arise from monochromatic signals which drift, or which experience a small amount of dispersion, or which have modulation with a very low index.

By "patterned signals" we mean those which possess some degree of organization or pattern which might be recognized as too well-structured to have come from a random process. Accordingly, one of the algorithms to be discussed later does search for clusters of signal characteristics, which can be viewed as a kind of pattern. In addition one of the primary reasons for introducing the human observer into the signal detection process is to take advantage of the capacity of the human to detect or recognize certain types of patterns.

## 2.7 - Review of Signal Processing Steps

As we saw in Section 2.3 there are three steps or stages in the process of searching for and understanding intelligent signals from space. These are:

1. Signal Detection
2. Signal Identification
3. Signal Decoding

Signal detection, the central focus of this study, is the process of distinguishing signals from background noise. In the next section we explore some theoretical considerations relating to this step. In Chapter 3 we present some specific signal detection algorithms implemented in this design.

Once we have detected a signal we turn to the problem of classifying it as: astronomical, RFI, or unidentified (including ETI). If a signal is identified as astronomical, it is sent to the observing astronomer for further study. If a signal is known to be RFI, it can be classified and cataloged, possibly to help with later identification of other RFI. If a signal cannot be identified as either astronomical or RFI, it is stored until such time as further identification efforts can be made. This recording or archiving can be done automatically or under operator discretion.

If a signal is believed to originate from an ETI, it must be interpreted or decoded by methods and algorithms yet to be devised.

## 2.8 - Signal Detection Algorithms

In this section we consider the general types of approaches or algorithms which can be applied to the problem of detecting signals in noise. Specific

algorithms selected for our designs are discussed in Chapter 3. We consider six approaches.

1. Data aggregation
2. Data addition
3. Search for sinusoids
4. Search for drifting sinusoids
5. Search for polarization
6. Pattern recognition

Data aggregation is a process of combining in any one of a number of ways the new data. A common reason for doing this is to reduce the amount of data which must be processed in a given amount of time. A second reason for aggregating data is to include in one value the effects of more than one signal property. For example, one term might be used to measure both power and polarization. If we wished only to pay attention to signals in which both power and polarizations exist, then a multiplicative aggregate term could be devised. A disadvantage of aggregation is that it sacrifices data. This effect may be overcome by storing the raw data for some time.

Data addition is a form of aggregation which usually is implemented with the special motivation of increasing the signal to noise ratio. Data values may be added over successive MCSA bins, in which case we call this coarse-binning. Data can also be added for the same bin or channel over a number of successive time records. The addition can also be performed over any other combinations of data values.

Data addition increases the signal to noise ratio by taking advantage of the fact that the data tends to add coherently while the noise adds incoherently. Data addition can be done over the amplitudes of the signals or over

the power (the sum of the squares of the real and imaginary components). It may be undesirable in some cases to add amplitudes, if for example these amplitudes are complex and the MCSA has introduced a phase shift which permits amplitudes to subtract.

There are a number of ways to search for sinusoids by taking advantage of their periodic nature. Most of these approaches are equivalent to narrowband filtering, as discussed in Section 2.5 above. One basic approach in our situation is to use the MCSA as a narrowband filter and then sum the output of each bin over many time records.

If the sinusoid is drifting (changing frequency slowly) we cannot add bin outputs since the signal will drift out of one bin into the next as time goes on. Two approaches are possible. The first is to force the center frequency of a channel or bin to track or follow a drifting input. Such a system would be in essence a phase-locked loop. A second approach is to add the signal over a set of channels which are different for each time period, and in this way follow or track the signal. The latter approach was implemented (see Chapter 3) in the Carrier Wave Detector.

The search for polarization must begin at the receiving antenna which has two probes, each of which is sensitive to one of two orthogonal polarizations. The receiver/MCSA system carries the two resulting signals along in parallel channels to the Signal Detector, where they are then used to obtain a measure of the degree of polarization of the input. In our implementation this data is used in three of the nine NBIT tests.

Pattern recognition is a very general term which can mean many different things. It refers to ways in which signals are grouped or organized in some sense. In the design which follows, the Cluster Detector is a form of pattern recognition device.

### References

1. Shklovskii, I.S. and Sagan, C. , Intelligent Life in the Universe, New York: Dell, 1967.
2. Sullivan, W.T. III, Brown, S., and Wetherill, C., "Eavesdropping: The Radio Signature of the Earth," Science, Vol. 199, January 27, 1978, pp.377-388.



## **Chapter 3**

### **System Algorithms**

Taking Three as the subject to reason about--  
A convenient number to state--  
We add Seven, and Ten, and then multiply out  
By One Thousand diminished by Eight.

The Result we proceed to divide, as you see,  
By Nine Hundred and Ninety and Two.  
Then subtract Seventeen and the answer must be  
Exactly and perfectly true.

The method employed I would gladly explain,  
while I have it so clear in my head,  
If I had but the time and you had but the brain  
But much remains yet to be said.

--Lewis Carroll, The Hunting of the Snark

### 3.1 - Purpose of This Chapter

This chapter will introduce the Oasis Signal Detector (OSD), explain its role in a search for intelligent signals, and develop, in some detail, the algorithms that this device employs as it shifts and sorts through gigabits of data.

As mentioned in Chapter 2, the OSD examines the output of the multi-channel spectrum analyzer, in order to recognize signals therein, and, subject to the control of the operator, determines the nature of the signals as one of

- a) RFI (radio frequency interference)

- b) natural astronomical bodies

- c) ETI (extraterrestrial transmissions).

The OSD also records interesting portions of the input data (selected by itself and/or the operator, at the operator's discretion), and at all times, monitors the entire SETI search system.

## 3.2 - Overview

### Real Time Processing

The OSD is a real time instrument, that is, it recognizes a signal, if any is present, during the time the antenna is pointed at a particular target, or shortly (16 minutes) thereafter. Once a signal is detected, identification tests are made with the aid of the operator and the OSD's RFI detection system, as well as the radio astronomy observer at his own output station, using a computerized catalog of radio sources. This rapid performance of identification tests eliminates costly and time-consuming reobservation of targets, and allows for the detection of transient types of signals of any of the three types. Reobservation will therefore occur only when there is a high probability that a detection was not a false alarm.

The length of observation of a target is nominally set at 1000 seconds (16.6 minutes), but since the desired length of time may vary appreciably, the system is configured to handle observations of any multiple of 20 seconds in length.

### Signal Recognition Procedure

The number of known signal modulation schemes is arbitrarily large. While it was felt important to design the OSD to be able to detect as many different signal forms as possible, we also felt that the OSD should be highly sensitive to those types of signals deemed optimal as interstellar beacons, as well as to high power signals leaked into space from the Earth. The most dominant components of these signals are very narrowband polarized carriers and pulsed RF signals, which to a distant observer may appear to be drifting slowly in frequency. To respond to these pulses and carrier waves, two special purpose

detectors were designed. They are called the pulse detector (PD) and the carrier wave detector (CWD). They are designed to be as sensitive as possible, with an entering data point integrated into many different sums and products, each of which represents a different type of possible drifting carrier or pulse.

Since artificially generated carrier waves are often constant in amplitude from second to second, relatively stationary in frequency, and always 100% polarized, a detection method highly sensitive to signals with such attributes was developed. This method (or measure), called Generalized Coherence, is described fully in the latter part of this chapter, and is employed within the CWD algorithm. Thus the OSD is highly alert to the presence of polarized carrier waves, even if they are Doppler-drifting.

Beyond providing signal-specific algorithms, the OSD has the facility to detect signals of an arbitrary and unknown nature. This is accomplished by a two-stage processor called the Numerical Battery of Independent Tests (NBIT). The first stage analyzes data from the MCSA in small parcels, and reports to the second stage a summary of the attributes of the data in each parcel. To this end, each parcel (called a block) undergoes a battery of tests. If any block scores outstandingly in any of the tests (which examine total power, polarization, and the presence of narrowband amplitude spikes, and more), the operator is alerted. In addition, a second stage always examines the report of the first stage (even in the absence of outstanding features). The second stage looks for the presence of trends, or patterns, such as the existence of blocks adjacent in time or frequency which possess similar attributes. This second stage is called the Cluster Detector (CD).

The terms PD, CWD, NBIT, and CD represent both the algorithms and the hardware processors that perform these operations. These four processors comprise the computational part of the OSD, called the realtime processing system, and are

described mathematically in this chapter, and as hardware devices in the next chapter. The realtime processing system is shown in Figure 3.1.

An "alarm" is the term used in this report to indicate an internal signal or flag from a processor to the human interface system describing the detection of a signal in noise. An alarm contains the information: What (how much), when, and where (frequency).

### Human Interface

The operator interface system consists of the operator, a central computer (which allows the operator the access to converse rapidly with the system), interactive graphic displays, a small archive in which either he or the central computer may decide to retain data segments of special interest, and a package of software routines which allow him to explore the data in his own way.

Most importantly, the operator is given access to the total observation archive which contains temporarily all of the data collected on the current target. He is shown the alarms as they come in, and is free to either let the system decide that which is important enough to save or else to make the decision himself. Providing aid in this task is the RFI detector, and the Radio Astronomy System, since each unit is capable of automatically informing the central computer that certain alarms are identifiable as non-ETI. Those signals of astrophysical interest will be processed and recorded independently by the Radio Astronomy System. The Human Interface System and the Radio Astronomy System are described in Chapters 5 and 6 respectively.

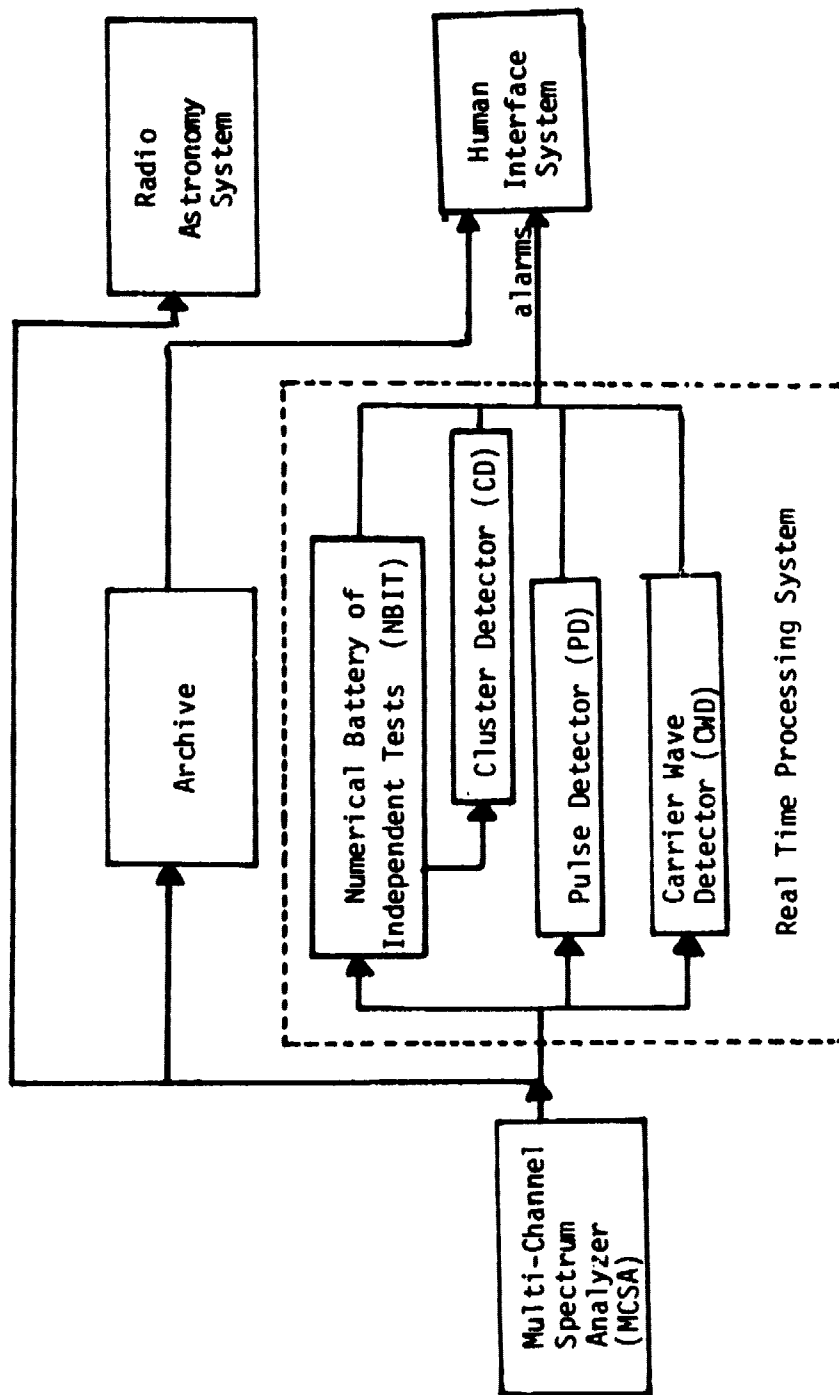


Figure 3.1 The Oasis Signal Detector

### 3.3 - Generalized Coherence

#### Background

It is widely believed that intelligent signals will be coherent with respect to both polarization and time. It is therefore desirable that the receiving system measure these two signal characteristics and use them to discriminate against signals which are not coherent, resulting in a decrease in sensitivity to incoherent signals. Since noise is incoherent, the discrimination against noise results in a substantial increase in sensitivity to coherent signals.

Polarization coherence and time coherence are two independent characteristics, and can be measured for any received signal. Polarization coherence is a measure of the consistency of the amplitude and phase between the signals received by two orthogonally polarized antennas. Time coherence is a measure of the consistency of the amplitude and phase of a single signal, with respect to a time-delayed version of itself. Although these two types of coherence are determined in quite different ways, it will be shown below that when both have been taken into account they lose their individual identities and merge into an overall Generalized Coherence.

## Polarization Matching

The receiving system contains two independent receivers, connected to orthogonally polarized feed antennas. Typically, the feed antennas will have horizontal and vertical, or left and right circular polarizations. It would be possible to process the data from each antenna separately, but this is undesirable because it essentially doubles the cost of data processing. A preferable procedure would be to combine the two antenna outputs by some method, and then process the resulting single signal, provided that this can be done without a significant decrease in sensitivity. One might even suspect that if the two antenna outputs were combined in some optimum way, the sensitivity would be further increased.

Measuring the Polarization of a Received Signal - In order to examine various methods of handling the two output signals, some basic polarization theory is required. Suppose the feed antennas are horizontally and vertically polarized, and denote their complex output voltages by  $x$  and  $y$ . This notation will be used throughout this report. (The choices of horizontal and vertical, as opposed to left and right circular, is arbitrary but generally done in polarization discussions. It will make no difference in the final result.) The type of polarization of the incoming wave is completely determined by the ratio of the magnitudes of the two outputs, and by their phase angle difference (Kraus, 1966, and Cohen, 1958) and is commonly expressed as two angles:

$$\delta = \angle x - \angle y \qquad -180^\circ \leq \delta \leq +180^\circ \qquad (3-1)$$

$$\gamma = \tan^{-1} \left| \frac{y}{x} \right| \qquad 0 \leq \gamma \leq 90^\circ \qquad (3-2)$$

which can be plotted on a spherical surface known as the Poincare sphere (see Figure 3.2).



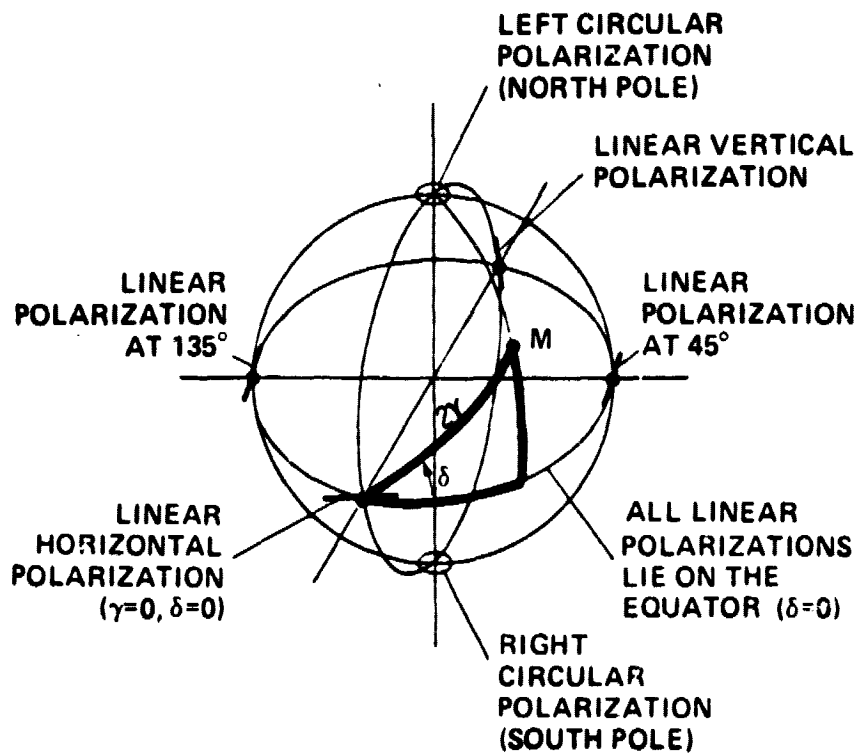


Figure 3.2 - The Poincaré Sphere

Having measured  $\delta$  and  $\gamma$  for any incoming signal, its polarization type may be plotted at the appropriate point M on the Poincaré sphere. For example, suppose the two outputs have equal magnitudes and differ in phase by  $90^\circ$ . Then  $\delta = 90^\circ$  and  $\gamma = \tan^{-1} 1 = 45^\circ$ . Plotting this on the Poincaré sphere, we see that the received signal is left circularly polarized.

Setting the Polarization of the Receiving Antenna to Match That of the Signal - Given two orthogonally polarized antennas, it is possible to combine their outputs in such a way as to synthesize a single antenna of any desired type of polarization. This is the inverse of the measurement problem, and may be stated in this way: Given the desired type of polarization, as specified by the angles  $\delta$  and  $\gamma$ , what are the weighting factors needed to combine the two antenna outputs in the form

$$Z = K_1 x + K_2 y \quad (3-3)$$

in such a way as to optimize reception for that type of polarization?

These weighting factors are derived in Appendix 3.1, giving the relation

$$Z = x \cos \gamma + y e^{j\delta} \sin \gamma \quad (3-4)$$

When searching for signals whose polarization type is unknown, a possible optimum method is to use the signal itself to control the x and y combining process, thereby adaptively matching the polarization type of the antenna to that of the incoming signal. This may be done by first measuring the type of polarization ( $\gamma$  and  $\delta$ ), and then synthesizing the antenna that matches that polarization. This combined operation can be carried out in real time, using the single relation (developed in Appendix A3.2),

$$Z = \left( \frac{|x|}{|x|^2 + |y|^2} \right) x + \left( \frac{|y|}{|x|^2 + |y|^2} \right) e^{j \left( \tan^{-1} \frac{\text{Im}(\overline{xy^*})}{\text{Re}(\overline{xy^*})} \right)} y \quad (3-5)$$

Comparison with Alternative Polarization-Handling Methods - One simple method for dealing with signal polarization is to add the powers coming out of the two antennas ignoring all polarization considerations. It is a fact that the total signal power will always be obtained by doing this, regardless of the signal or receiving antenna polarization. However, the fact that two independent noise powers are added always decreases the SNR by a factor of  $\sqrt{2}$ , where the polarization matching method weights the two so that the loss ranges from 0 to  $\sqrt{2}$ . Another disadvantage is that the signal cannot be tested for time coherence at a later stage in the analysis, if power is calculated at this point.

A second simple method is to add the amplitudes. This corresponds to setting  $K_1=K_2=1$  in equation (3-3) above and means in reality that we are receiving at a single polarization (linear at  $45^\circ$ , assuming x and y antennas). Any other polarization will suffer some loss, ranging up to infinite for the opposite case (linear at  $135^\circ$ , for x and y antennas).

The two polarizations could also be processed completely independently. This means that we now are receiving with two polarizations (corresponding to whatever type of antennas are being used). Other polarizations suffer some loss, up to 3 db for those halfway between the two in use. (e.g. - if x and y antennas are used, then left and right hand circular both have 3 db loss.)

By combining the antenna outputs in various ways (i.e., choosing  $K_1$  and  $K_2$ ), any number of specific receiving polarizations can be synthesized. In Project Cyclops it was suggested that 4 or 6 cardinaly chosen polarizations be created in this way, and processed separately. This reduces the possible loss, since a signal polarization must then fall closer on the average to one of the receiver polarizations. The disadvantage of this is that the data processing requirements increase in direct proportion to the number of polarizations.

All of these various methods are summarized in Table 3.1. The average losses were calculated assuming the signal polarization types are uniformly distributed around the Poincare sphere.

<u>Method</u>	<u>Number of Polarizations Searched</u>	<u>Minimum Loss (db)</u>	<u>Maximum Loss (db)</u>	<u>Average Loss (db)</u>	<u>Relative Amount of Data Processing Required</u>
Add Powers of the 2 Antennas	0	1.5	1.5	1.5	1
Add Amplitudes of the 2 Antennas	1	0	$\infty$	5.1	1
Process 2 Antennas Separately	2	0	3.0	1.3	2
Synthesize 4 Polarizations and Process Separately	4	0	3.0	0.7	4
Synthesize 6 Polarizations and Process Separately	6	0	1.0	0.4	6
Signal-Matching	$\infty$	0	0	0	1

Table 3.1 - Comparison of different polarization processing methods

It is important that the polarization synthesis not be done in the RF or IF portions of the receivers, as is commonly done. To do so means that all signals at any instant of time will be received with the same polarization. On the other hand, if the polarization synthesis is done as part of the data processing (as in the signal-matching method), then each signal can be received with a different polarization.

### Polarization Coherence

Partial Polarization - It is not necessary for real-world signals to be polarized at all. Most natural radio sources radiate unpolarized signals. Some natural radio sources are partially polarized, radiating signals containing both polarized and unpolarized components. Any signal radiated from a single antenna, however, is always completely polarized. We use this fact to great advantage.

The power received by an antenna of polarization type M in response to a signal of polarization type M' is

$$P_r = P_s \left[ 1/2 + m \left( \cos^2 \frac{MM'}{2} - 1/2 \right) \right] \quad (3-6)$$

where  $P_s$  is the signal power

$m$  is the degree of polarization;  $0 \leq m \leq 1$

$MM'$  is the polarization mismatch angle between the signal and antenna,

as plotted on the Poincare sphere;  $0 \leq MM' \leq \pi$

When the signal is completely unpolarized (i.e.,  $m = 0$ ), the received power is one-half that of the signal power, regardless of what receiving antenna polarization is used. When the signal is completely polarized (i.e.,  $m = 1$ ),

the received power lies between zero and the signal power, depending on the polarization mismatch angle.

The degree of polarization is defined in Appendix 3.2 as

$$m = \sqrt{1 - \frac{4(\overline{xx^*} \overline{yy^*} - \overline{xy^*} \overline{x^*y})}{(\overline{xx^*} + \overline{yy^*})^2}}$$

The degree of polarization may be regarded as the radius of the Poincare sphere. A completely polarized signal has radius of one (i.e., the unit sphere), whereas a completely unpolarized signal has a radius of zero (i.e., a dot at the center). This graphical representation serves to illustrate the fact that the type of polarization (location on the surface of the Poincare sphere) becomes less and less important as the degree of polarization (radius of the Poincare sphere) becomes smaller and smaller. In the extreme case of an unpolarized signal, the type of polarization ceases to exist, as exemplified by the remaining dot at the center.

Polarized Signal Component Extraction - When the receiving antenna polarization is matched to that of the signal (e.g., using the polarization matching method described above), equation (3-6) becomes,

$$P_r = P_s \left( \frac{1}{2} + \frac{m}{2} \right).$$

This may be rewritten as

$$P_r = P_s \left( \frac{1-m}{2} + m \right). \quad (3-7)$$

Here the first term in the brackets is the unpolarized component of the signal and the second term is the polarized component. Since only half of the unpolarized component is received, a polarization-matched antenna acts as a filter, leaving the polarized component unchanged. If the powers in

two antennas were simply added, all of the unpolarized power would be received thereby decreasing the SNR by a factor of 2, neglecting receiver noise.

Since we are seeking signals which are polarized, only the second term is of interest. We wish to extract that fraction of the received power that is polarized, i.e.,

$$P_{rp} = P_r \left( \frac{mP_s}{P_s \left( \frac{1-m}{2} + m \right)} \right)$$

$$P_{rp} = P_r \left( \frac{2m}{1+m} \right) \quad (3-8)$$

If we were concerned only with signal power at this point we would just compute the desired result according to equation (3-8). To do so, however, would prevent later examination of the time coherence properties of the signal. Instead, we multiply the complex signal  $z$  of equation (3-5) by the square root of the fraction given in (3-8) (since  $z$  is an amplitude, as opposed to power). This extracts the polarized component of the composite signal  $z$ , i.e.,

$$z_p = z \sqrt{\frac{2m}{1+m}} \quad (3-9)$$

The effects of noise on this calculation are given in Appendix A3.3.

An alternative to extracting the polarized signal component is to ignore the degree of polarization. The disadvantage of this is that no discrimination is made against unpolarized signals. Two signals of equal power, one polarized and the other unpolarized, would give the same receiver output. If the degree of polarization is used as described above, the polarized signal would be unaffected, whereas the unpolarized signal would be eliminated from the receiver output.

## Time Coherence

The complex degree of self-coherence of a signal  $z(t)$  is defined as

$$\mu(\tau) = \frac{\overline{z(t+\tau) z(t)^*}}{\sqrt{\overline{z(t) z(t)^*}} \sqrt{\overline{z(t+\tau) z(t+\tau)^*}}}$$

This is in fact the same as the complex normalized autocorrelation function of the signal. The magnitude of this quantity is known as the Degree of Self Coherence, and  $0 \leq |\mu(\tau)| \leq 1$ . The denominator of  $\mu(\tau)$  is the power of the signal, so that

$$|\mu(\tau)| = \frac{|\overline{z(t+\tau) z(t)^*}|}{P_z}$$

It should be noted that the magnitude of  $\mu$  depends on both the phase and amplitude of the original signal. It is not correct to say that the magnitude of  $\mu$  depends only on the amplitude of the original signal, or that the phase angle of  $\mu$  depends only on the phase angle of the signal. Thus  $|\mu|$  is a measure of both amplitude coherence and phase coherence in the signal.

We are interested in only that part of the received signal and noise combination that has time coherence. That component may be extracted by multiplying the received power by the degree of self-coherence, i.e.,

$$P_c = \text{Coherent Power} = |\mu(\tau)| P_z = \left| \overline{z(t+\tau) z(t)^*} \right|$$

The time lag  $\tau$  used above may be regarded as the coherence time of the power measurement. It must be sufficiently long that the noise component shows no coherence, but no longer than necessary, to allow for slow signal variations. The MCSA provides independent outputs at one



second intervals, so  $\tau$  must be an integral number of seconds. A lag of zero will cause the noise to appear coherent, and is therefore too small. A lag of one or more will make the noise independent from point to point. Therefore we choose a lag of one, as being long enough to discriminate against noise, yet short enough to allow slow signal variations. Thus,

$$P_C = \text{Coherent Power} = \left| \overline{z(t+1) z(t)^*} \right| \quad (3-10)$$

Coherent power has the useful property that it is unaffected by any phase shifts that are linear functions of time (see Appendix A3.4). This is crucial for the current application because the MCSA introduces a linear phase shift into the signal that is a function of the unknown position of the signal frequency within a frequency bin. Another way of saying this is that the complex output voltage of an MCSA channel that contains a non-drifting sinusoidal signal, is another sinusoidal signal. The frequency of that output signal is the difference between the center frequency of the bin and the original signal frequency. Thus for 1 Hz bins, the output frequency lies in the range of  $\pm 1/2$  Hz. Then if the amplitude of this sinusoid is  $A$ , we have from Appendix A3.4 that

$$P_C = \frac{1}{N-1} \left| \sum_{i=1}^{N-1} A_i A_{i+1} \right| = A^2 = P_S ,$$

so that for a sinusoidal signal, the coherent power is identical to the signal power.

Consider now an incoherent signal (such as a natural radio source), whose power is equal to that of the above sinusoidal signal. The MCSA output will be a voltage whose amplitude and phase are random variables, and whose spectrum is band limited ( $\pm 1/2$  Hz) white noise. The coherent power will be the same

as the signal power for the first data point, and will decrease as  $\frac{1}{\sqrt{N}}$  as additional data points are included in the measurement. Thus coherent power has an advantage over signal power in that it distinguishes between coherent and incoherent signals.

### Generalized Coherence

Formulation - In the preceding sections, specific operations were carried out on the received signal, taking into account successively the type of polarization, the degree of polarization, and the time coherence. These operations are shown schematically in Figure 3.3. If we now combine these three operations into a single expression, and label the resulting quantity as Generalized Coherence Value (GCV), we obtain (conceptually) (see Appendix A3.5 for details)

$$GCV = F \left\{ P_{xn} R_{xn}(1) + P_{xyn} [R_{yx}(1) e^{-j\delta} + R_{xy}(1) e^{+j\delta}] + P_{yn} R_{yy}(1) \right\}$$

where

$$\delta = \tan^{-1} \frac{\text{Im}(R_{xy}(0))}{\text{Re}(R_{xy}(0))}$$

$$F = \frac{\sqrt{[R_{xx}(0) - R_{yy}(0)]^2 + 4 R_{xy}(0)^2}}{R_{xx}(0) + R_{yy}(0)}$$

$P_{in}$  is the normalized power in antenna  $i$

$P_{ijn}$  is the normalized cross power between antennas  $i$  and  $j$

$R_{ij}$  is the correlation function between the  $i^{\text{th}}$  and  $j^{\text{th}}$  antenna outputs.

In essence, all the possible auto- and cross-correlation functions, both time shifted and unshifted, appear in the GCV expression, in a symmetrical way.

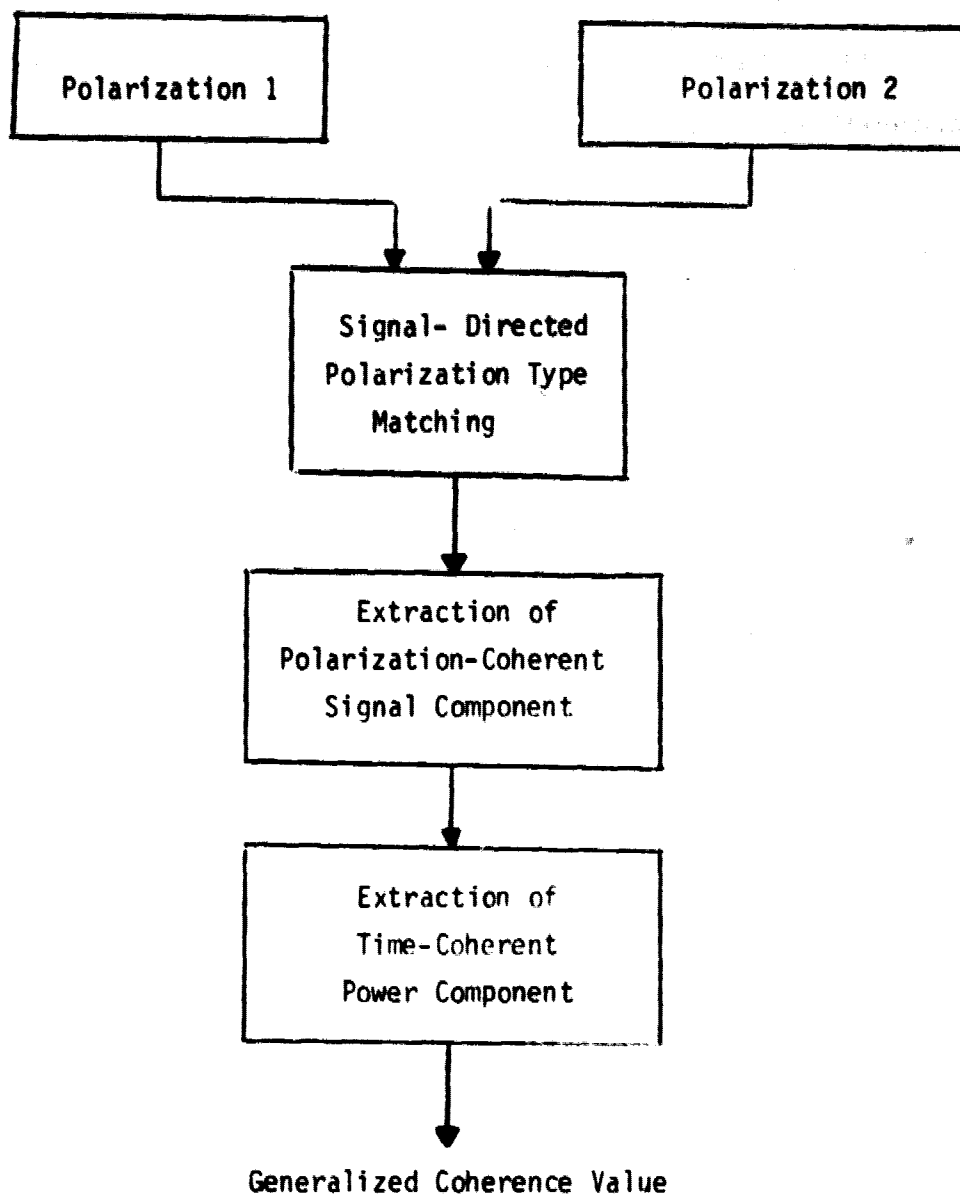


Figure 3.3 - Operations Involved in Calculation of the Generalized Coherence Value

It has been demonstrated previously (Born and Wolfe, 1970) that the coherence of a signal obeys the electromagnetic wave equation, and propagates through space in the same manner as the signal itself. This fact lends additional support to the idea of coherence being in some sense a fundamental quality of the signal.

Real-Time Evaluation - Calculation of generalized coherence involves calculating the correlation functions described above. All these quantities are summations of combinations of the instantaneous outputs of the two antennas. In a realtime computing situation, the summations can be calculated continuously as the data points become available, providing the partial results of the autocorrelation function are accumulated. When the desired amount of data has been accumulated, the summations may be combined by relatively simple computations to give the generalized coherence value.

When very many outputs (channels) are monitored for generalized coherence concurrently, the retention of the partial results may require substantial overhead in terms of storage. An alternative scheme developed in Chapter 4 enables the serial evaluation of channel outputs, and thus eliminates the need for partial result accumulations.

Performance - Noise causes the computed values of the correlation functions described above to have statistical fluctuations. In addition, if the correlation functions are calculated using too few data points, they will be biased toward the high side (as are all correlation functions).

In order to quantify the statistical characteristics of GCV, computer simulations were run for a number of signal-to-noise ratios.

Figure 3.4 is a typical example of the simulations made to investigate GCV. The data from all such simulations is gathered together in Figure 3.5. Details are given in Appendix A3.6.

It is possible that GCV should be squared for optimum detection. This is suggested by the fact that its probability density is similar to that of a sinusoidal signal in noise, just prior to square law detection in a normal receiver, and because GCV is fundamentally the magnitude of a vector even though it has units of power. GCV is the sum of many vectors, analogously to the way that a traditional coherent detector adds up many vectors prior to squaring their sum.

Inspection of Figure 3.5 reveals that below a SNR of 0.6, the power calculation is more sensitive, and above 0.6 GCV is more sensitive, and becomes increasingly more so as SNR increases. This advantage will be manifested as higher detection probabilities and lower false alarm rates. The small amount of data shown for the square of GCV reveals an even greater advantage. The power sensitivity crossover is then 0.4, and the output SNR curve climbs more rapidly. Actually, the relative performance of GCV vs. power is significantly better than the curves indicate since the power curves assume the signal polarization is known beforehand and exactly matched to a single chosen antenna, with the other antenna and its accompanying receiver noise being ignored. The GCV curves, in contrast, assume no prior knowledge of the signal polarization and include the effects of noise from both receivers. It is believed that future studies comparing randomly mismatched total power detection schemes to GCV detection will show GCV the favored approach to even lower SNR signals.

This can be summarized by saying that the generalized coherence technique achieves its goal of using the signal polarization type information in a useful way, and of discriminating against unpolarized and incoherent signals, without significantly sacrificing sensitivity, and in fact provides a significant enhancement to sensitivity.

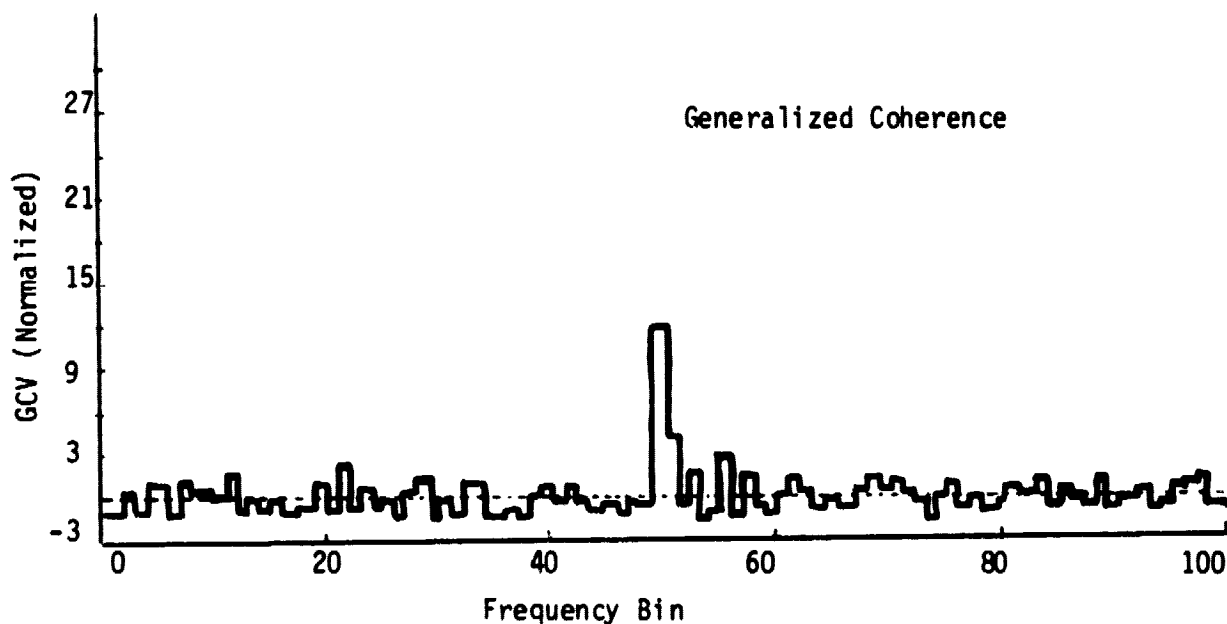
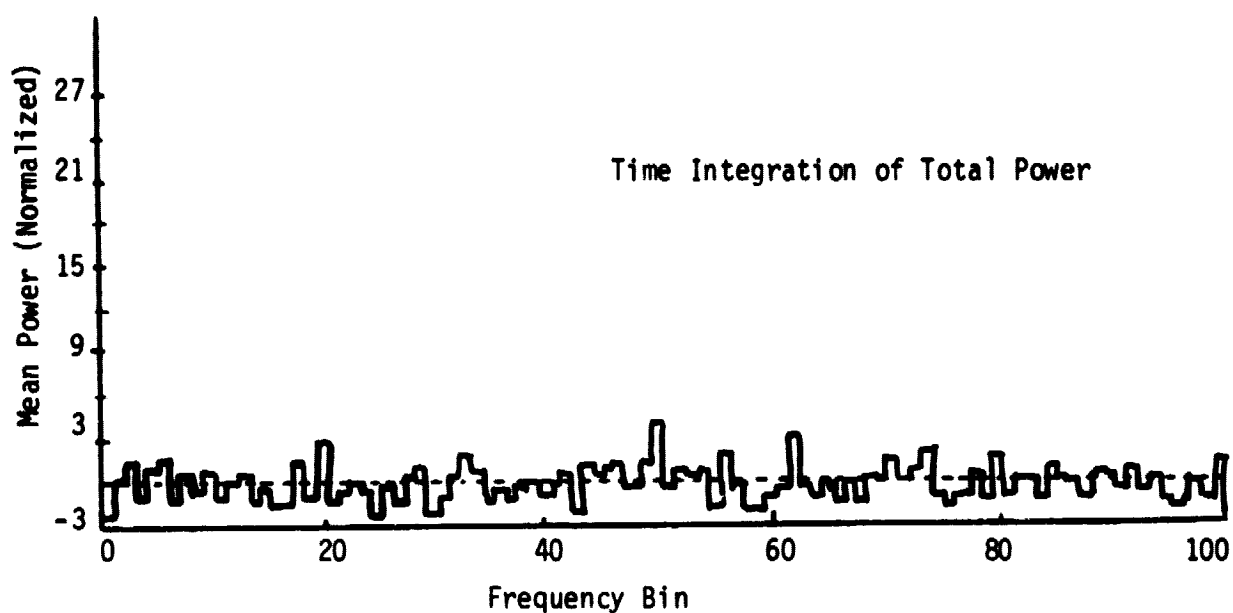


Figure 3.4 - Comparison of Time Integration of Total Power with Generalized Coherence, with a Signal to Noise Ratio of 1.0 (in a Single Matched Polarization), calculated for 25 Successive Data Points.

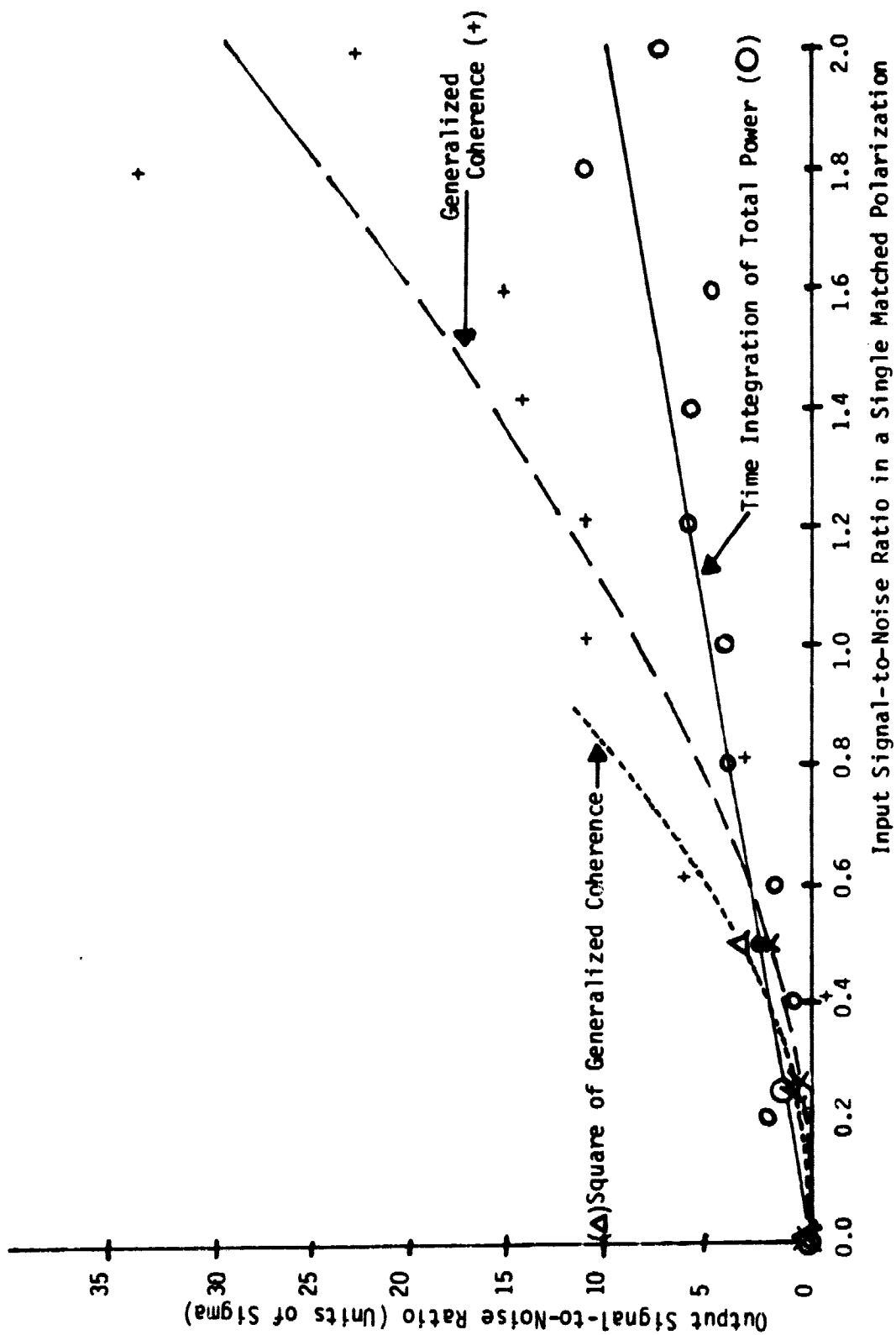


Figure 3.5 - Comparison of Detection Methods

### 3.4 - Analysis of Variance (ANOVA) Algorithm

#### Background

A shortcoming of most detection algorithms is the degree to which they have been designed to detect specific types of signals. While this philosophy serves to maximize sensitivity to these particular signals, many are blind to even strong signals outside this range. One algorithm that demands a minimum of prior assumptions is an analysis of variance (ANOVA). On the time-frequency matrix of the MCSA's output, ANOVA can perform three tests of the null hypothesis (that noise alone is present) to search for significant differences between means of rows, columns, and interaction terms. If the null hypothesis is not once rejected, ANOVA concludes that no signal is present. But if rejected even once, not only has ANOVA addressed the primary question of signal detection, but valuable information about the type of signal present is provided--with only a simple thresholding of the three results of a single standardized calculation.

For example, if a monochromatic signal is present, it will manifest itself in the ANOVA matrix by increasing the power in the column that corresponds to its frequency. The presence of this type of signal would be detected in the ANOVA by a rejection of the null hypothesis of column effects. Reflecting the difference between the column means associated with the signal and those column means associated only with noise, the value for the sum of squares of column sums would increase, resulting in an increased variance ratio for columns. If the variance ratio for columns exceeds a present critical variance ratio, the null hypothesis is rejected.



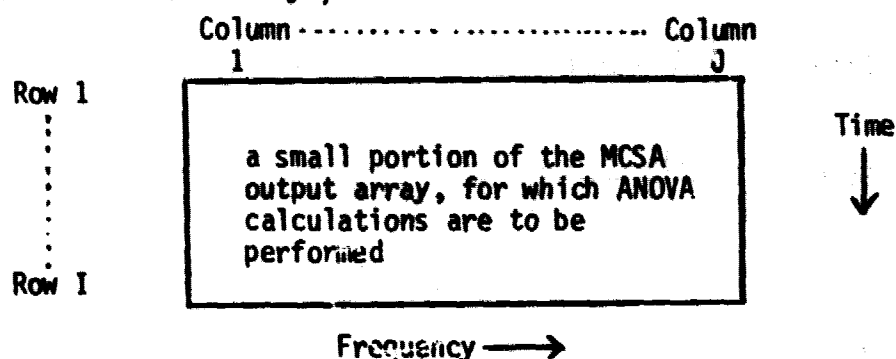
Similarly, a broad band signal would increase the power in all the matrix columns within the signal's bandwidth. This would result in higher column means for those columns containing the signal, producing a proportionately more significant variance ratio for the columns.

Row means and the associated row variance ratios are affected by pulsed broadband signals which distribute power across many frequencies within one second. As with the column variance ratio, the effect on the row variance ratio of more than one pulse within the matrix is cumulative, the optimum algorithm response occurring when the signal fills one row of the cells of the ANOVA matrix.

Rapidly drifting signals are detected primarily by the interaction term of ANOVA, though significant row and column effects can result as well, depending upon the slope of the path of signal as it moves through the matrix. Drift rates greater than one Hz/sec being detected in rows, while rates that are less than one Hz/sec are detected by the column hypothesis.

Because ANOVA makes no assumptions about the signal to be received and demands only standardized calculations, it serves as an important global signal detector for the SETI problem. Though further analysis is required after detection, the results of the three tests of null hypothesis (Row, Column, and Interaction) greatly reduce the problem of signal type identification.

The concepts of ANOVA calculations will be illustrated here by word equations and examples, and the detailed mathematical theory is explained in Appendix A3.8. We will assume here that the time-frequency array output from the MCSA has a zero mean (i.e., that the noise goes equally above and below zero, as will be shown is provided by baseline removal, a standard procedure of radio astronomy.)



The three ANOVA calculations are:

$$\text{ROW EFFECT} = \frac{\left( \text{average of all points in Row 1} \right)^2 + \left( \text{average of all points in Row 2} \right)^2 + \dots + \left( \text{average of all points in Row J} \right)^2}{I \cdot (\text{average of the square of all points in the array})}$$

$$\text{COLUMN EFFECT} = \frac{\left( \text{average of all points in Column 1} \right)^2 + \left( \text{average of all points in Col.2} \right)^2 + \dots + \left( \text{average of all points in Col.J} \right)^2}{J \cdot (\text{average of the square of all the points in the array})}$$

$$\text{INTERACTION EFFECT} = \frac{\sum_{\text{all } i,j} \left( \text{point}_{i,j} - \frac{\text{average of all points in Row } i}{J} - \frac{\text{average of all points in Col. } j}{I} \right)^2}{I \cdot J \cdot (\text{average of the square of all points in the array})}$$

All three calculations involve the averages of averages, and are power ratios. All are normalized to the average power.

The following examples illustrate the sensitivities of the three calculations to various types of data structure. For simplicity, data values are limited to  $\pm 1$  and 0. Note that the number of +1 values equals the number of -1

values, so that the zero-mean assumption is valid.

Example 1:

1 1 1  
0 0 0  
-1 -1 -1

$$\text{Row Effect} = \frac{(1)^2 + (0)^2 + (-1)^2}{3 (6/9)} = 1$$

$$\text{Column Effect} = \frac{(0)^2 + (0)^2 + (0)^2}{3 (6/9)} = 0$$

$$\begin{aligned} \text{Interaction Effect} = & \frac{(1-1-0)^2 + (1-1-0)^2 + (1-1-0)^2 + (0-0-0)^2 + (0-0-0)^2}{3 \cdot 3 \cdot (6/9)} + \\ & \frac{(0-0-0)^2 + (-1+1-0)^2 + (-1+1-0)^2 + (-1+1-0)^2}{3 \cdot 3 \cdot (6/9)} = 0 \end{aligned}$$

This example illustrates how the Row Effect detects long horizontal data structures.

Example 2:

1 0 -1  
1 0 -1  
1 0 -1

$$\text{Row Effect} = \frac{(0)^2 + (0)^2 + (0)^2}{3 \cdot (6/9)} = 0$$

$$\text{Column Effect} = \frac{(1)^2 + (0)^2 + (-1)^2}{3 \cdot (6/9)} = 1$$

$$\begin{aligned} \text{Interaction Effect} = & \frac{(1-0-1)^2 + (0-0-0)^2 + (-1-0+1)^2 + (1-0-1)^2 + (0-0-0)^2}{3 \cdot 3 \cdot (6/9)} + \\ & \frac{(-1-0+1)^2 + (1-0-1)^2 + (0-0-0)^2 + (-1-0+1)^2}{3 \cdot 3 \cdot (6/9)} = 0 \end{aligned}$$

This example illustrates how the Column Effect detects long vertical data structures.

Example 3:

$$\begin{array}{ccc} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{array}$$

$$\text{Row Effect} = \frac{(0)^2 + (1/3)^2 + (-1/3)^2}{3 \cdot (6/9)} = \frac{1}{9}$$

$$\text{Column Effect} = \frac{(0)^2 + (0)^2 + (0)^2}{3 \cdot (6/9)} = 0$$

$$\begin{aligned} \text{Interaction Effect} = & \frac{(1-0-0)^2 + (0-0-0)^2 + (-1-0-0)^2 + (0-1/3-0)^2}{3 \cdot 3 \cdot (6/9)} + \\ & \frac{(1-1/3-0)^2 + (0-1/3-0)^2 + (-1+1/3-0)^2}{3 \cdot 3 \cdot (6/9)} + \\ & \frac{(-1+1/3-0)^2 + (1+1/3-0)^2}{3 \cdot 3 \cdot (6/9)} = \frac{8}{9} \end{aligned}$$

This example illustrates how the Interaction Effect detects long diagonal data structures. Note that although all the samples use square arrays, this is of course not necessary for ANOVA calculations. These examples have also used infinitely high SNR's--to clearly illustrate the calculation. Evaluation of the technique was, of course, conducted on low SNR signals, and the detection sensitivity was thus explored as shown in Appendix A3.9.

## Performance

The potential of Analysis of Variance as a signal detector has been evaluated by means of computer simulations in a series of systematic experiments. The computer simulation involves modeling the noise field as a  $(10 \times 40 \times 2)$  matrix of exponentially distributed random variables. 40 represents the number of frequency channels (columns), 10 represents the number of independent time samples (rows), and 2 represents the number of observations (data values) per cell. This size matrix is that used in the NBIT algorithm for reasons to be described later in this chapter. At least 2 values per cell are necessary for the calculation of the interaction term in ANOVA. This term is important for the detection of certain signal types and for the reduction in the size of the residual variance. Since the residual variance is the denominator of the variance ratio which is used for tests of hypothesis, a reduction in it increases the overall sensitivity of detection. Although ideally, replication should be obtained from the simultaneous observation of the event of interest, the simulations are carried out in the more realistic fashion of letting the first row consist of data taken during the first and second seconds of observation and each subsequent pair of scans fill the remaining rows. Thus a total of 20 independent time samples are used. Superimposed upon the noise field is a signal of given specification.

The data, obtained in this fashion, are then analyzed by ANOVA technique. The end product of the ANOVA algorithm is the calculation of the variance ratio for rows, columns, and interactions. When the variance ratio exceeds a preset critical value or threshold, the null hypothesis is rejected and the signal is detected. By examining the values of the variance

ratio under systematically varying conditions, an understanding of ANOVA's effectiveness as a detector unfolds.

The simulation experiments and detailed results are described in Appendix A3.9.

### Conclusions

The simulation studies indicate the possible strengths and weaknesses of ANOVA as a signal detector. For the matrix dimensions presently being considered, it seems ideally suited to the detection of a broadband pulse. Narrowband non-drifting and slowly drifting signals are also handled well. The one real failure occurs with the rapidly drifting signal (slope -1). It is notable that changing the matrix dimensions to a square would certainly improve ANOVA's capability to detect this type of signal, while not degrading the performance of detecting the previously mentioned types.

Based on the simulation results involving bit quantization, a reasonable bit size for the data processing using ANOVA has been determined to be  $N = 8$ , as is used in the implementation. Thus, in all, ANOVA presents itself as a significant global signal detector for the SETI problem. No assumptions need be made about the type of signal. Once one is detected, further analysis is in order. By retaining information on which of the three hypotheses (Row, Columns, and Interactions) are significant, the problem of identification of signal type is significantly reduced.

### 3.5 - Oasis System Algorithms

#### Overview

The Oasis Signal Detector is a Real-Time Processing System containing four basic components: the Pulse Detector (PD), the Carrier Wave Detector (CWD)

which is meant to detect narrowband carrier signals that may be drifting slowly in frequency, the Numerical Battery of Independent Tests (NBIT), which consists of a collection of statistical tests followed by a pattern recognition algorithm, the Cluster Detector (CD), that is designed to detect non-random associations among the signal parameters in N-space. This system is shown conceptually in Figure 3.6.

### Data Stream Preparation

We are assuming in this study that two of the functions performed by the MCSA are:

- Quantization level preparation
- Baseline removal

Data within the MCSA is clipped at times both by truncating low order and high order bits. This process has been given much consideration by the designers of the MCSA via simulations, and, that the quantization levels be renormalized linearly, is a requirement of the MCSA implementation. Based on the simulation studies it is felt that each quantization level corresponds to .3σ of the signal power. Implementation design for this is given in Chapter 4.

Baseline removal is accomplished within the MCSA through the use of its low resolution 57 kHz output channel. A region of the sky slightly offset from the target to record the broadband spectral responses of the receiver,  $R(f)$ ; and this function is then used to normalize true spectra of the target  $T(f)$ . Normalization is accomplished in the following manner. The coarse resolution values are used to compute

$$R'(f) = \frac{R(f) - T(f)}{T(f)}$$

yielding a zero mean normalized (except for local power excesses) spectrum,  $R'(f)$ .

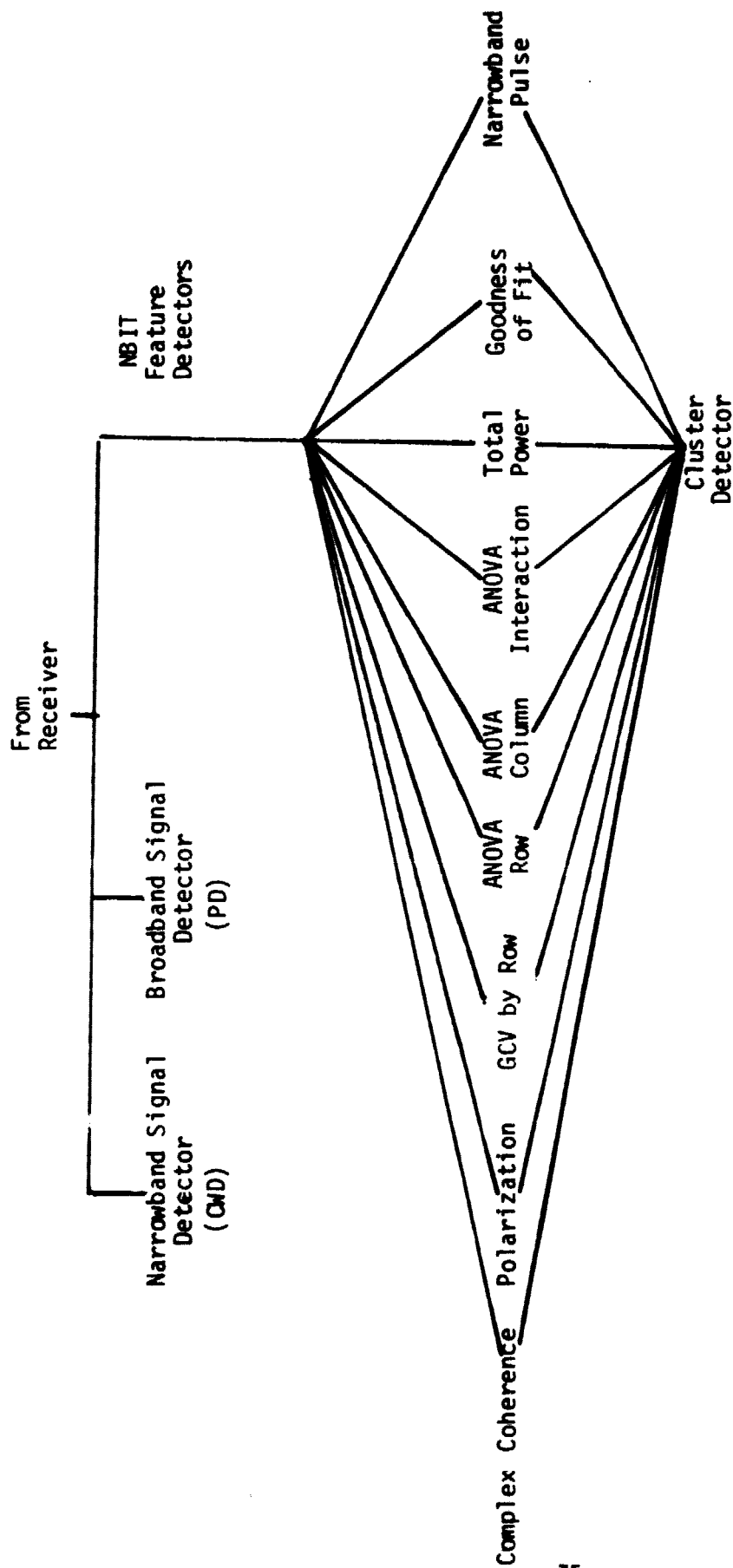


Figure 3.6 - OASIS System Signal Detectors



This practice is common and essential in Radio Astronomy to remove receiver gain effects before observation of their source.

### Pulse Detector (PD)

The coarse binning stage of the system examines each spectral scan under a wide range of resolutions, to search for excess total power that might signify a signal of interest. Although it can be argued that very narrowband signals are optimal for interstellar communication, to neglect the possibility of detecting broadband signals or pulses within the data, when such information is easily accessible, would be a mistake. Such a wideband signal may occur in the eavesdropping or the pulse-beacon cases of signal transmission.

Under the most ideal condition, when the bin width corresponds to a transmitted signal's bandwidth, coarse binning serves as a matched signal for signal detection. However, since the signal bandwidth and central frequency are not known, a wide range of bandwidths is used. The algorithm first sums the  $8 \times 10^6$  adjacent frequency bins pairwise in non-overlapping pairs to produce  $4 \times 10^6$  channels times 2 Hz/channel resolutions which are compared to a threshold. This procedure is repeated again and again until every set of  $2^k$  ( $0 \leq k \leq 23$ ) adjacent bins has been combined and tested. Any coarse bin which has a total power above the predetermined threshold causes an alarm to be sent to the central computer. It should be noted that one scan (one second of MCSA output) can be analyzed at all 23 resolutions within one second. Simultaneously, an integrated scan, the sum of all previous 8 MHz x 1 Hz resolution scans is similarly binned and tested for excesses due to repeatedly detected pulses.

## Carrier Wave Detector (CWD)

Overview - The Carrier Wave Detector (CWD) is designed to search for narrow band signals which may be drifting slowly in frequency due to the Doppler shifts. The CWD works on the 2-dimensional array, having dimensions of frequency and time, in which a drifting carrier would appear as a straight line whose angle of inclination with respect to the vertical is less than  $45^{\circ}$ , i.e., a drift rate  $-1 \text{ Hz/sec} \leq r \leq 1 \text{ Hz/sec}$ .

In principle, the CWD evaluates the data by overlaying the array with a "mask" which consists of 20 rays, each of which has a different slope. The rays all intersect at a single time-frequency element of the array and form a double fan (Figure 3.7). A column of 20 consecutively observed spectral elements (20 seconds of data arranged one below the other) is passed under the mask so that each of the 8 million single Hertz channels acts as the center point of the mask exactly once. For each ray of the mask, the CWD weights each  $k^{\text{th}}$  row entry (Figure 3.8) of the incoming data for the way in which the ray intercepts the array, and then a Generalized Coherence Value (GCV) is computed which reflects the intensity and coherence of the weighted data overlayed by the ray. The results from each new 20 seconds of data analysis are tested against a threshold. If any GCV exceeds the threshold, the detector sends an alarm to the central processor. In addition, all GCV's are accumulated for the entire 1000 seconds of observation time. At the end of the total observation time, the accumulated GCV's are also compared to a threshold. These accumulated values are also evaluated with the coarse binning technique, to search for extended drifting features.

In order to take advantage of the fact that polarization and phase information are provided by the pair of MCSA's, the CWD operates on the

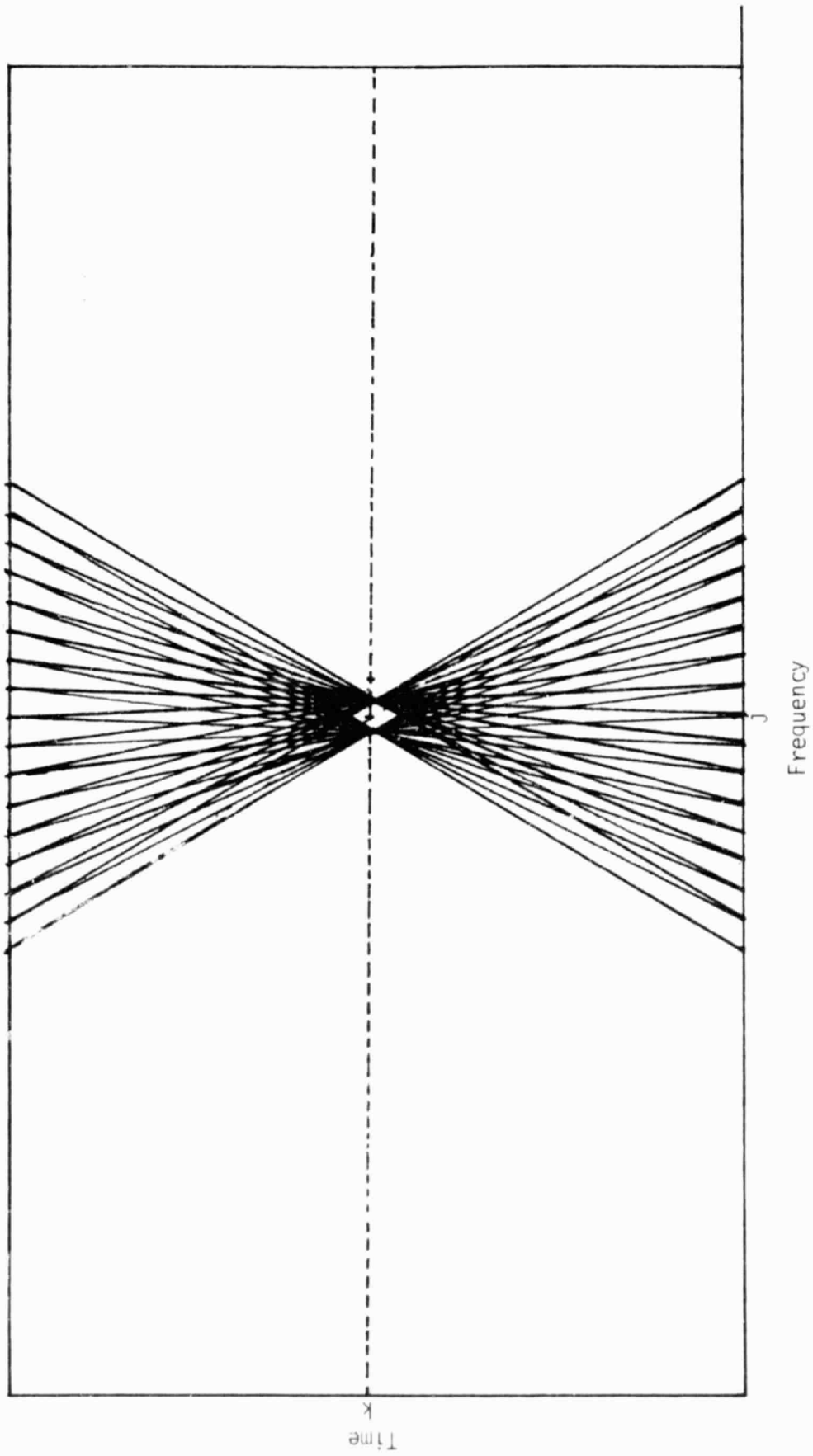


Figure 3. 7 - Two DCD Time-Frequency Masks. Both are Centered at Time  $k$ , and they are at Frequencies  $j$  and  $j-1$ .

Generalized Coherence Values (GCV's) described above. For each angle and each center frequency, one GCV is computed using 20 values, each of which is a weighted entry for that ray (Figure 3.8). The weighting process involves both amplitude and phase, and is discussed in detail in Appendix A3.10.

### Numerical Battery of Independent Tests (NBIT)

Introduction - Since there are numerous possibilities for the forms of extra-terrestrial intelligent signals, the signal detection problem can be approached in two ways: (a) a set of tests each of which has a very high probability of detecting signals of a very specific type; or (b) a set of tests that have a lesser probability of recognizing a much more general class of signals.

The Pulse Detector, which searches for pulses of various widths, and the Carrier Wave Detector, which searches for narrowband signals at various drift rates, are examples of the first approach. However, because the characteristics of an ETI signal are purely conjectural, to search even a significant portion of the possibilities within the general class would require a prohibitive amount of memory and/or processing.

The second approach is exemplified by the numerical battery of independent tests (NBIT). The NBIT signal detection system is a two stage technique which reduces the data to a set of 9-bit values in the first stage, and employs a pattern recognition algorithm on the compacted array in the second stage to detect signals.

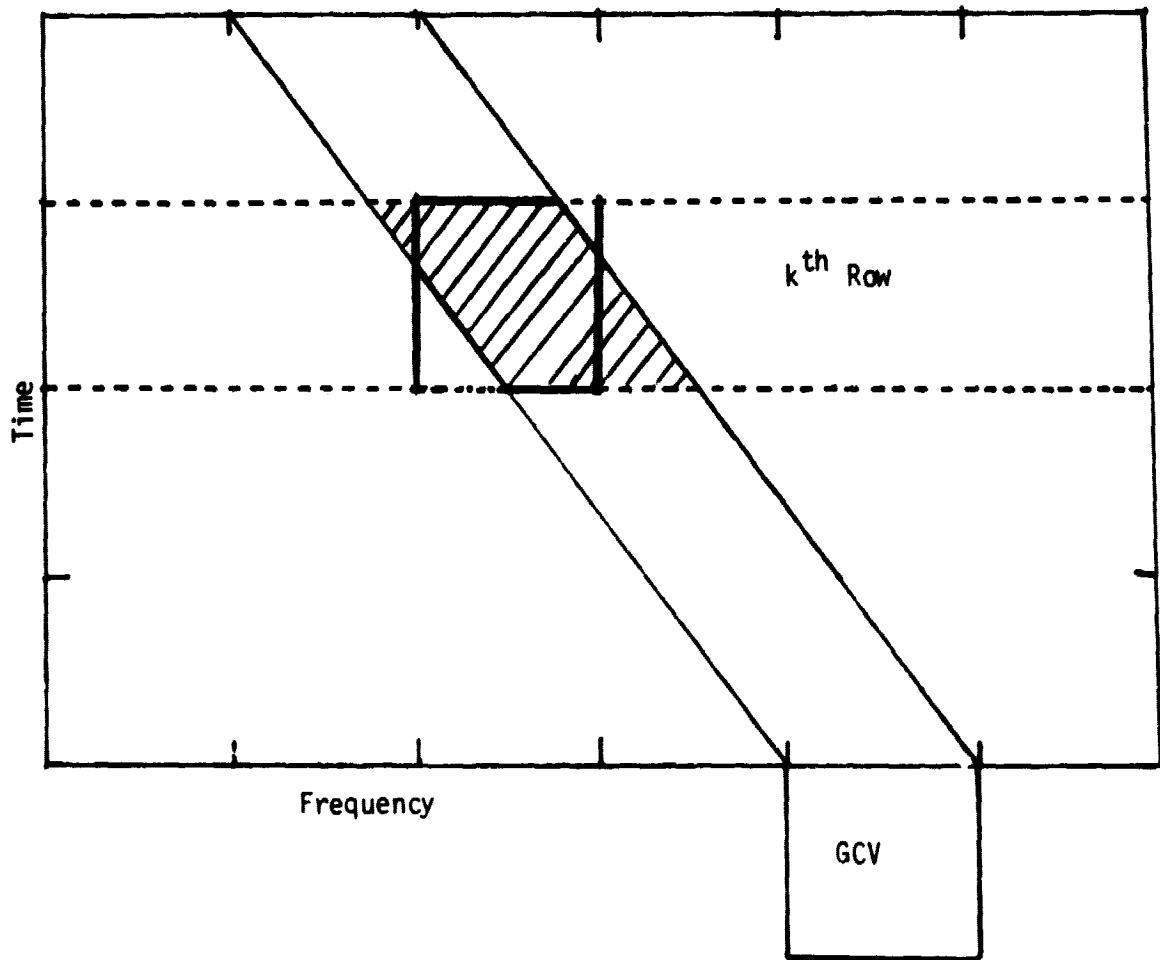


Figure 3.8 - Weighted Contribution to GCV of Each Time-Frequency Measurement

In the first stage the entire array is subdivided into blocks, each of which is 40 Hz by 20 seconds in frequency-time space. Each block has an 8 Hz overlap with adjacent blocks in order to ameliorate the aliasing problem between blocks. Nine statistical hypothesis tests are applied to each block of incoming data. One bit of the 9-bit number is assigned for each test, and the bit is set when the null hypothesis is rejected. The first stage tests include: total power in the two polarizations; degree of polarization; complex amplitude coherence detection; broadband coherent pulse detection for each polarization using generalized coherent values; 8-hertz pulse detection for each polarization; ANOVA analysis of variance for rows (frequency), ANOVA columns (time) and ANOVA interaction, for each polarization; and goodness of fit for each polarization.

For each of the tests, the null hypothesis is rejected at one of the two thresholds ( $\alpha$ -levels). The first level (low-level) threshold is set to trigger relatively easily ( $\alpha = .5$ , which means that 50% of the time the bit for a particular test under conditions of noise alone will be "on"). The second level (high-level) threshold, on the other hand, corresponds to a very low  $\alpha$ -level in order to ensure a very low false alarm rate for immediate central computer processing (the threshold values are adjusted to reflect the results of simulation or actual operation). The first level threshold is set so that the difference between the false alarm rate under noise ( $\alpha$ ) and the probability of detection ( $1-\beta$ ) for the signal type for which the test is designed is maximized for each algorithm. This criterion is developed further in Chapter 5.

The 9-bit number for each block in the first stage is stored in a "pattern block" which is passed on to a two-step cluster detection processor. The first step of the processor tabulates a histogram on the occurrence of various 9-bit values and sets an alarm for the central computer if a statistically high number of specific "pattern points" occurs. The second step, a single linkage cluster seeking algorithm, searches the entire pattern block for subtle signal patterns and sets an alarm if a pattern emerges.

The two-step cluster detection algorithm which analyzes the array is a non-deterministic signal detection scheme, and as such has a great deal of flexibility.

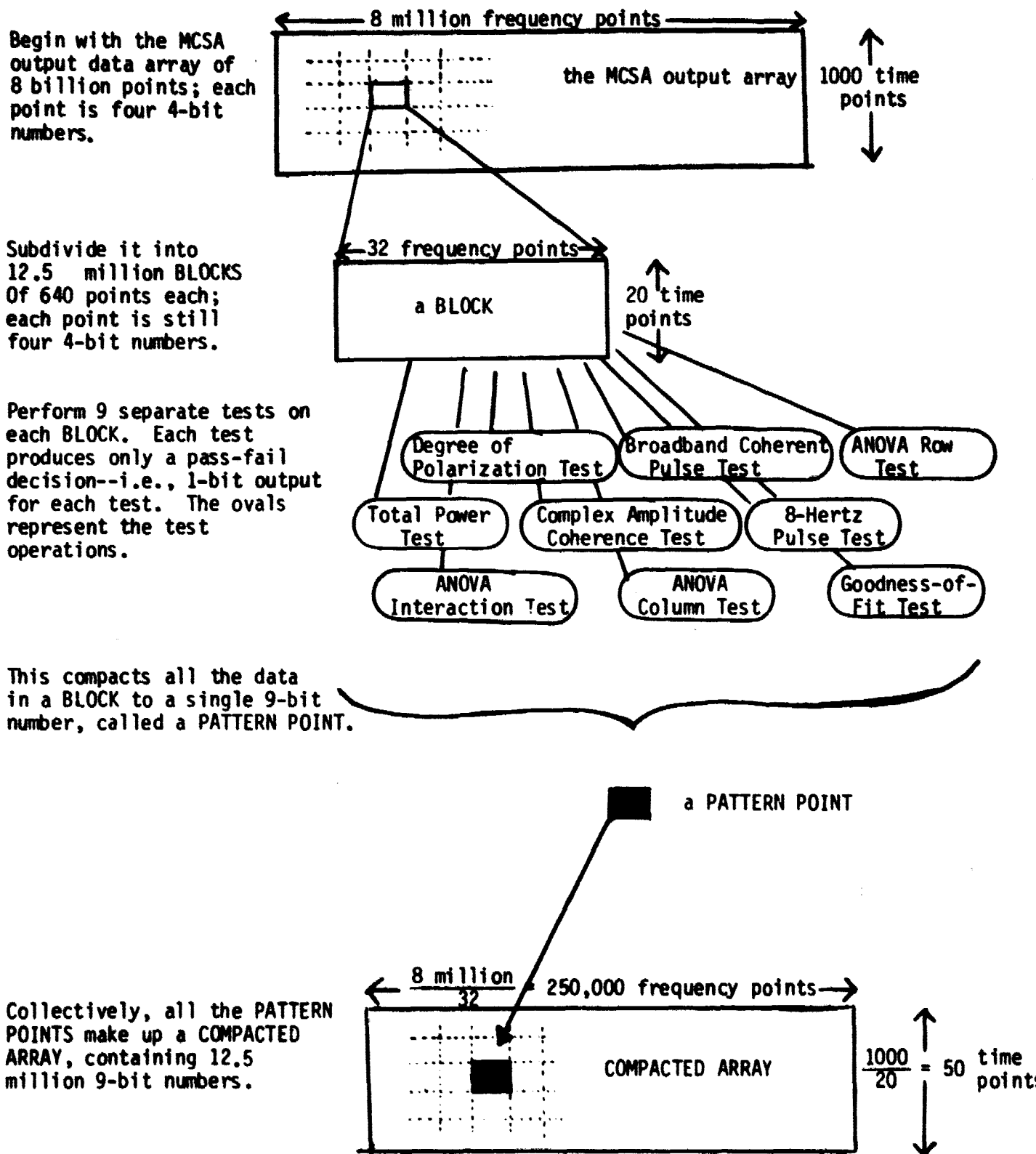
In general the entire NBIT system has several advantages over more traditional signal detection systems. The two-stage nature of the system allows a "second look" at the data in real time and it employs two independent and distinct methods. The meaningful compaction of the data in the first stage makes it possible to store information for the entire search duration at a reasonable cost, which is vitally necessary to detect a weak signal of prolonged duration. The NBIT signal detection system is also versatile because it is not signal-type specific.

Figure 3.9 provides an overview of the NBIT and Cluster Detection Algorithms.

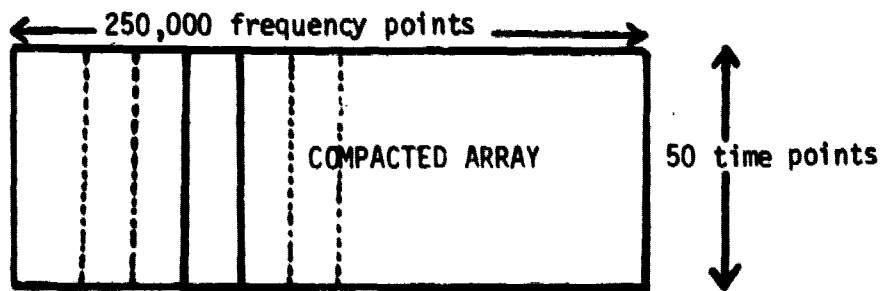
For the sake of conceptual clarity, Figure 3.9 intentionally omits the following details:

1. A BLOCK overlaps its neighbors by 4 points on each side, so its width is actually 40 points, and total content is 800 points. All other quantitative numbers are correct as shown.

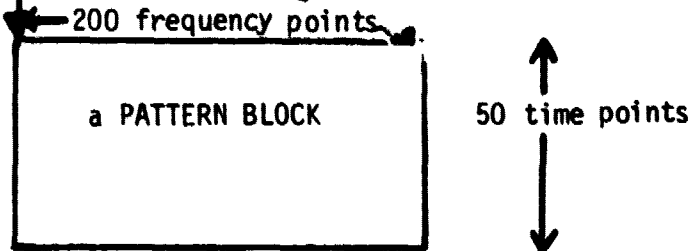
**Figure 3.9**  
**An Overview of the NBIT and Cluster Detection Algorithms**



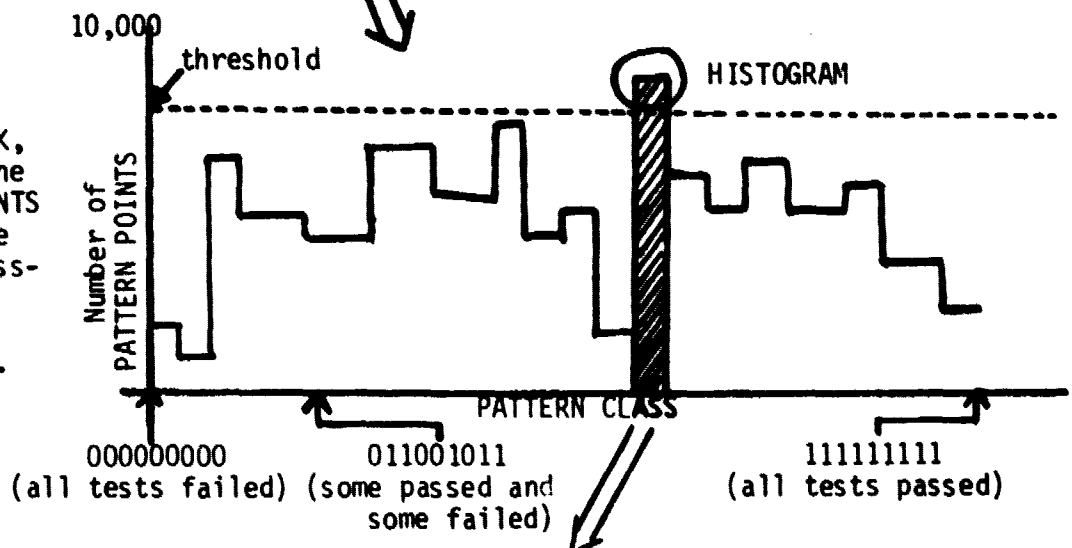




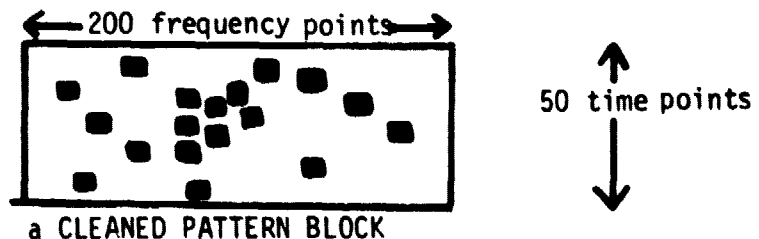
Divide the COMPACTED ARRAY into 1250 PATTERN BLOCKS of 10,000 PATTERN POINTS each. Each point is still a 9-bit number. The division allows for computer processing of each pattern block.



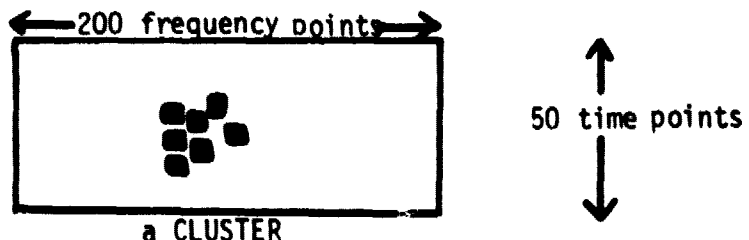
For each PATTERN BLOCK, make a HISTOGRAM of the number of PATTERN POINTS in which each possible combination of the pass-fail tests occurred. There are  $2^9$  possible combinations of tests. Each combination is called a PATTERN CLASS.



For any PATTERN CLASS that exceeds the significance threshold on the HISTOGRAM, make a CLEANED PATTERN BLOCK, that maps only the occurrences of the corresponding combinations of tests.



Search the CLEANED PATTERN BLOCK for CLUSTERING, using an adjacency algorithm. Remove all PATTERN POINTS that are not part of a CLUSTER. If the number of remaining PATTERN POINTS is significant, decide that a signal has been found.



Note: Some features have been simplified for clarity on this diagram. See the text for actual details.

2. There are additional signal detection decisions made at two intermediate levels:
  - a. The 9 tests each have two thresholds. If the high-level threshold is exceeded, a detection is immediately indicated. The low-level threshold is set such that it is exceeded by noise 50% of the time, and is used to create the 9-bit PATTERN POINT.
  - b. If any HISTOGRAM has an extremely high point, a detection is immediately indicated, regardless of the presence or absence of clustering.
3. The 9 tests are actually histogrammed with regard to on bits "1" and don't care bits "\*", not the on and off "0" bits. This allows for additional useful test combinations like total power and degree of polarization and ignores everything else. Thus there are actually  $2^9 = 512$  PATTERN CLASSES.
4. HISTOGRAM points are not actually chosen for cluster searching on the basis of their absolute level. Rather, the available data processing power and time are optimized by choosing the highest HISTOGRAM bin excesses globally over all PATTERN BLOCKS.
5. The cluster search algorithm detects not only "blobs" as illustrated, but also finds and follows long linear features.

## The Individual NBIT Tests

1. Total Power - The purpose of this test is to determine whether the total power accumulated over a 20-second by 40-Hz block is significantly different from that expected from white noise alone. As the input to the MCSA is Fourier Transformed, the power of any pulse or wide-band signal is spread across the frequency domain. Summing the power over a range of frequencies and time may accumulate the power to a detectable level.

$$\text{Total Power} = P_T = \sum_{i=1}^{800} (|x_i|^2 + |y_i|^2)$$

This value is then compared with previously determined first and second level threshold values.

If this power exceeds either of the two pre-determined thresholds, a bit is set and sent to the CPU for further processing. Because the white noise is thought to be fairly constant, and this test sets a bit approximately fifty percent of the time, the first power threshold will only be slightly above the noise level ( $1\sigma$ ). The second threshold, if exceeded, immediately notifies the operator, and all information from that scan is recorded. This threshold is adjusted as discussed in Chapter 5.

The type of signal that is most easily detected by this method is a wideband signal. The rationale for this argument is that the algorithm is an integration technique and since the threshold is rather low, any amount of signal in the 40 Hz block will cause the 1-bit to be set. Since most wide-band signals are translated into the frequency domain as a set of sinc

functions, a relatively large number of points in the block will contain the signal. When these points are summed together the power in the block will be above the threshold level. Narrowband signals on the other hand, appear as approximate delta functions, and as such, the signal will be in relatively few points in the block. If the power of the signal is low enough, a narrowband signal may go undetected by this test. However, this algorithm is not specifically sensitive to any signal type, but rather the elevation in power, over the noise level, that comes with having a signal present in the block.

2. Degree of Polarization - A signal emanating from any single antenna will be completely polarized. Even when degraded by noise, a sufficiently strong signal will have a resulting degree of polarization significantly greater than that of noise alone.

In this test degree-of-polarization is calculated along each row of the 20 sec x 40 Hz block, and compared to a threshold value. Degree-of-polarization is calculated along rows in order to maximize sensitivity to pulses and other signal types that lack the time coherence required for detection by the CWD.

The results of this test can also be applied to the signal recognition problem. Signals of high power but low degree of polarization are unlikely to be ETI signals, while a high degree of polarization is a hallmark of artificial signals.

Procedure - A general discussion of degree of polarization can be found in Section 3.3. For the purposes of this test, the following

calculation is performed for each row of the current block:

$$m = \frac{1}{|\bar{x}|^2 + |\bar{y}|^2} \sqrt{(|\bar{x}|^2 - |\bar{y}|^2)^2 + 4 |\overline{xy^*}|^2}$$

where  $|\bar{x}|^2 = \frac{1}{40} \sum_{i=1}^{40} x_i x_i^*$ ,  $|\bar{y}|^2 = \frac{1}{40} \sum_{i=1}^{40} y_i y_i^*$

and  $\overline{xy^*} = \frac{1}{40} \sum_{i=1}^{40} x_i y_i^*$ . Each calculated value of  $m$  is compared to stored thresholds.

Sensitivity - If an overall false alarm rate of  $\alpha_0$  is desired for each block, then per row the false alarm rate is  $\alpha = 1 - (1 - \alpha_0)^{1/20}$ . (Then if  $\alpha_0 = .5$ ,  $\alpha = .034$ .)

The behavior of the degree of polarization under noise was determined empirically. For an integration length of 40, noise has mean degree of polarization .176 and standard deviation .078 (based on a sample size of 200). This is apparently independent of noise power.

If  $\alpha_0 = .5$  is chosen, then the threshold value for bit setting is .32, either by determination of the 93rd percentile of the data or by assumption of normality.

Signals of various strengths and durations were simulated, and their degrees of polarization calculated. Figure 3.10 summarizes the results, with each point representing an average of 50 values. (Note: signal width is the number of bins in a given row actually containing the signal, even if they are not contiguous.)

Examination of Figure 3.10 indicates that if degree of polarization  $m$  results from width = A and SNR = B, then for any d the same value of  $m$

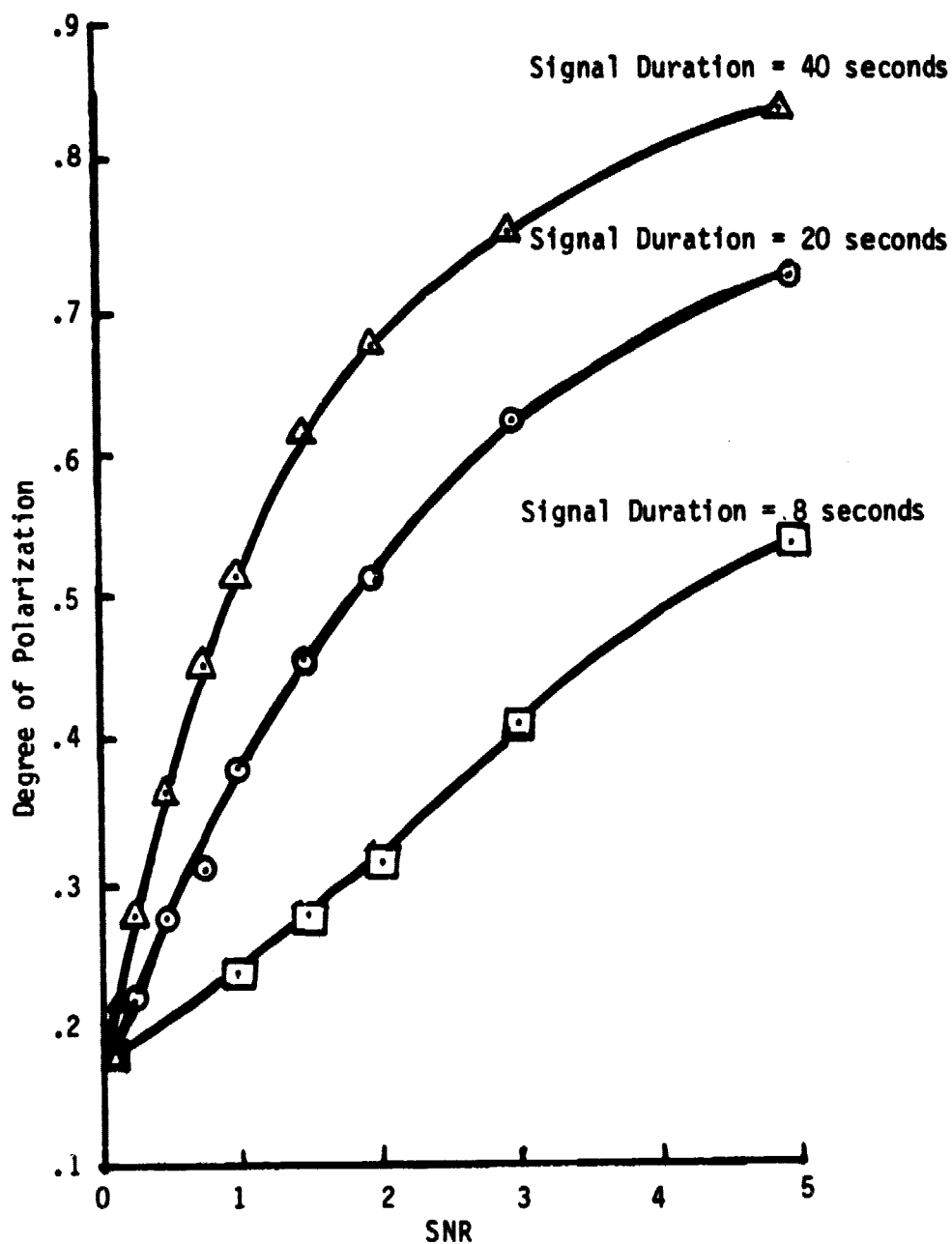


Figure 3.10 - Sensitivity of the Degree of Polarization Test to Signals of Varying SNR, for Various Signal Time Durations

**Number of Rows Containing a Signal of Bandwidth 40 Hz**

<u>SNR</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	<u>8</u>	<u>9</u>	<u>10</u>
.16	.12	.23	.32	.40	.47	.54	.59	.64	.68	.72
.25	.34	.56.	.71	.81	.87	.92	.95	.96	.98	.98
.50	.72	.92	.98	.98	1.00	1.00	1.00	1.00	1.00	1.00
.75	.93	.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1	.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
>1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.001	1.00	1.00

**Table 3.2 - Probability of Signal Detection by the Degree of Polarization Test, as a Function of Signal Time Duration (number of rows) and SNR**

will result from width =  $dA$  and  $SNR = B/d$ . In other words, degree of polarization is proportional to  $\frac{SNR}{\text{signal width}}$ .

Table 3.2 indicates detection probabilities for a signal of width 40 occurring in  $r$  rows of the block. This table can be applied to signals of width less than 40 by using the appropriately factored SNR. (Note: There is close agreement between the probabilities obtained from the samples and those derived from normality assumptions.)

3. Complex Amplitude Coherence - This test is designed to detect any polarized signal which may be continuous, pulsed, sporadic, or amorphous. The amplitudes need not be particularly large, as we are looking for persistent coherent signals.

For each 40 Hz by 20 sec. block of data, we compute

$$|A| = \left| \frac{\overline{xy^*}}{\sqrt{\overline{x} \overline{y}}} \right|$$

where  $\overline{xy^*}$  refers to the average cross power in the two MCSA's,  $\overline{x}$  and  $\overline{y}$  refer to the power in the x- and y-MCSA respectively. The magnitude of  $|A|$  is an indication of the correlation (or coherence) between the two polarizations; for example, if the two outputs are linearly related,  $|A| = 1$ , while if there is no correlation whatever between the two,  $|A| = 0$ . At the end of processing for each block  $|A|$  is compared with the predetermined first and second level thresholds.



#### 4. Generalized Coherence by Row (A Broadband Coherent Pulse Detector)

Wideband pulses will appear in a string of contiguous frequency bins simultaneously, for only one time period. They may then repeat at some later time. Such a signal could be generated at the transmitter by a sinc time variation. This would cause the received signal phase to have a linearly increasing variation with frequency, at a slope inversely proportional to the bandwidth of the pulse.

As shown in Appendix A2.1, the MCSA does not introduce any more than a linear relative phase shift between adjacent frequency bins. Therefore, the original linear phase variation with frequency is maintained at the MCSA output. The generalized coherence value may be computed across frequency rows, since it is insensitive to linear phase variations.

Specifically, the GCV is calculated across all 40 frequency bins, separately for each row. If the GCV for any row exceeds the predetermined threshold(s), the appropriate bit(s) for this test will be set.

5. 8-Hertz Pulse Detector - This test provides a means for detecting infrequently occurring pulses whose widths are a few Hertz. It computes the average power  $P$  in an 8-Hz span across each frequency row for each polarization.

$$P_x = \frac{1}{8} \sum_{i=0}^7 |x_i|^2$$

$$P_y = \frac{1}{8} \sum_{i=0}^7 |y_i|^2$$

where  $x_i$  and  $y_i$  are row outputs for the x-and y-MCSA's respectively. If any of these results exceeds the first level threshold, the 1-bit is set. Likewise, if the results surpass the second level threshold, such an alarm is set as well.

6,7,8. Analysis of Variance (ANOVA) - ANOVA can be used to detect several different types of signals. For example, assume that a narrowband monochromatic signal is present. This signal will manifest itself in the ANOVA matrix by increasing the power in the column that corresponds to its frequency. The presence of this type of signal would be detected in the ANOVA by a rejection of the null hypothesis of column effects. That is, the column mean associated with the signal would be higher than the column means associated with the noise. This difference in mean values is reflected accordingly in the column sum-of-squares, which in turn results in an increased F-ratio for columns. If the F-ratio for columns exceeds a preset critical F-value, the null hypothesis is rejected. Similarly, a broadband signal would increase the power in all the matrix cells associated with its frequency bandwidth. This again would result in an increase of column totals for those columns containing the signal, and ANOVA would pick this up by producing a significant F-ratio for the columns. In the event that a broadband continuous signal completely covers the number of columns in the ANOVA, this would be detected by comparing the total power (which is calculated as one of the steps in ANOVA) with the total power in the previously established noise matrix. Thus, narrowband and broadband continuous signals would be detected by a rejection of the hypothesis associated with columns.

A pulsed signal which is broadband and periodic would result in higher power reading in the row totals (for those rows where it appeared) and would result in a rejection of the hypothesis associated with row effects. It is apparent that the signal to noise ratio is a critical

parameter in any detection scheme. Simulation studies of SNR and ANOVA sensitivity are discussed above in Section 3.4.

A drifting signal (or a broadband signal that has higher frequencies arriving at regularly spaced intervals of time) would be detected primarily by the interaction term of ANOVA. Significant row and column effects can result as well, depending upon the slope of the path of the signal as it moves through the matrix. Slopes greater than minus 1 are detected by the row hypothesis. A tall spiked signal that appears somewhere in the matrix would be detected by row and column hypothesis and possibly even the interaction.

The ANOVA tests are conducted for each polarization using the power contained in each bin of the 20 second by 40 Hz block. The F-ratios for each of the three tests are compared with the thresholds and the appropriate action taken.

9. Goodness of Fit -This test compares the actual probability distribution of the MCSA output, treated as a quantized variable, with its expected distribution in the presence of Gaussian white noise. It is a particularly useful test since it makes no assumptions about the nature of the signal to be detected.

For this test the complex output of the MCSA is treated as a pair of real numbers for each 1 Hz frequency bin every second. Each block contains 800 pairs, so, for the two polarizations, there are a total of 1600 pairs. The set of output values for each member of the pair is compared with a Gaussian distribution, representing the expected outcome for white noise alone: For any possible output value  $k$ , let  $T_k$  be the theoretically

expected number of occurrences of  $k$  and let  $N_k$  be the observed number of occurrences.

Then 
$$T = \sum_k \frac{(N_k - T_k)^2}{T_k}$$
 is nearly a  $\chi^2$  (chi-squared)

random variable with  $M - 1$  degrees of freedom, where  $M$  is the number of possible output values for  $k$ . This value of  $T$  is compared against stored thresholds.

### Pattern Recognition - The Stage II NBIT Algorithm

1. Introduction - A pattern is simply the description of an object. Pattern recognition is considered one of the most common of activities of man, as well as all other living organisms. For extraterrestrial life to survive, it too must be able to produce and recognize patterns in order to communicate about and react to the dangers inherent in its environment. Once life--be it terrestrial or extraterrestrial--develops communication capabilities, the signals used for communication must contain patterns distinguishable from randomness. At the interstellar communication level, energy radiated across interstellar space must have, if nothing else, distinctive patterns that can be recognized as an artifact against the background of natural radiation. It is a primary function of the Oasis Signal Detector to recognize these patterns.

In the past, every SETI experiment began with the assumption that the hypothetical extraterrestrial radio signal would take some particular form. The appropriate signal analysis and detection scheme was implemented, the experiment was run, and the negative finding could be expressed in a single sentence: "No signal of the particular form and exceeding  $x$  flux density was found."

These restricted searches had the advantage of being relatively easy to mount, but their narrowly defined negative results were never very satisfying. Designing the signal detector as a mathematical pattern recognition system broadens the constraints on the search immeasurably. In particular, a mathematical pattern recognition system can take the parallel outputs of a number of analytic procedures applied to incoming signals and make a single decision based on the location of the point in feature space which they represent.

2. General Considerations - Since the MCSA provides a multidimensional output (frequency, amplitude, polarization, phase, and time), a generalized pattern recognition algorithm has been deemed a more useful tool than any pattern-specific signal search method, such as a matched filter. Furthermore, given the fact that planetary rotation would cause a transmitted pattern to drift across frequencies at an unpredictable rate, the need for a general spatial search for patterns is further accentuated. The pattern clearly visible in Figure 2.3a can not be recognized in any frequency scan. In the figure imagine frequency increasing to the right and time downward. If we added signal power vertically in time, the signal would again be lost in noise. The only cue as to its artificiality is its spatial structure in frequency-time space.

Although the human brain with eyes and ears as its input receivers, is the most powerful and versatile pattern recognition system we know of and could easily recognize a pattern such as the onesshown in Figure 2.3 , it is also subject to fatigue and hallucinations, it is selective and imprecise, and for

our purposes, too costly. The pattern or spatial non-homogeneity shown in Figure 3.11 (a) could be interpreted in many ways by the mind, or not be recognized as a "pattern" but instead be classified as a random phenomenon such as shown in Figure 3.11 (b).



Figure 3.11

Furthermore, the enormous data rate provided by the MCSA eliminates the human operator from consideration as a preliminary realtime pattern recognizer. If the raw data were to be displayed in real-time and in standard TV display densities, the latest 1000 seconds of accumulated spectra would require a console 43 miles long! (See section 4.1)

Powerful algorithms exist for recognizing well-defined prototype signal categories and for clustering (grouping) unknown categories sharing similar properties (Andenburg, 1973, Fu, 1976, and Ton and Gonzales, 1974). Some of these algorithms have been modified to suit different objectives in a detection and recognition scheme. In view of the implementation limits set by research time available in this study, we developed the simple yet extremely general algorithm described in this section. We feel that more sophisticated algorithms could easily be applied to the problem, and suggest study of their implementation.

3. Definitions - For sake of clarity and preciseness in the ensuing discussion we need to make the following definitions:

- a) A scan refers to one spectrum - 8 MHz x 1 Hz, with 1.6 bits per channel.
- b) The entire array refers to 1000 seconds of the MCSA output and is a 1000 sec. x 8 MHz matrix of raw data.
- c) A block is a 20 sec. x 40 Hz section of the entire array.
- d) The compacted array is a 50 x 250000 matrix of 9-bit numbers.
- e) A pattern block is a 50 x 200 matrix of 9-bit numbers which is a vertical strip of the compacted array.
- f) A pattern class is a set of 9-bit numbers with specified bits equal to 1 which determine the class and no constraint on the unspecified bits. Thus, both (0, 1, 1, 1, 0, 1, 0, 1, 1) and (1, 1, 0, 1, 1, 1, 0, 1, 0) belong to the class (\*, 1, \*, 1, \*, \*, \*, 1, \*), and, also, to the class (\*, 1, \*, \*, \*, \*, \*, \*, \*), while (0, 0, 1, 1, 0, 1, 0, 1, 1) belongs to neither of these classes.
- g) A pattern point is any 9-bit number of the compacted array.
- h) A cell is the time-frequency location of a pattern point. Note the one-to-one correspondence between blocks and cells.

Figure 3.12 should further clarify the above definitions.

4. The Rationale for Cluster Analysis - Cluster analysis is a pattern recognition technique which encompasses many diverse methods for discovering structure within complex bodies of data. The common characteristic of the algorithms of cluster analysis is the attempted grouping of data units into

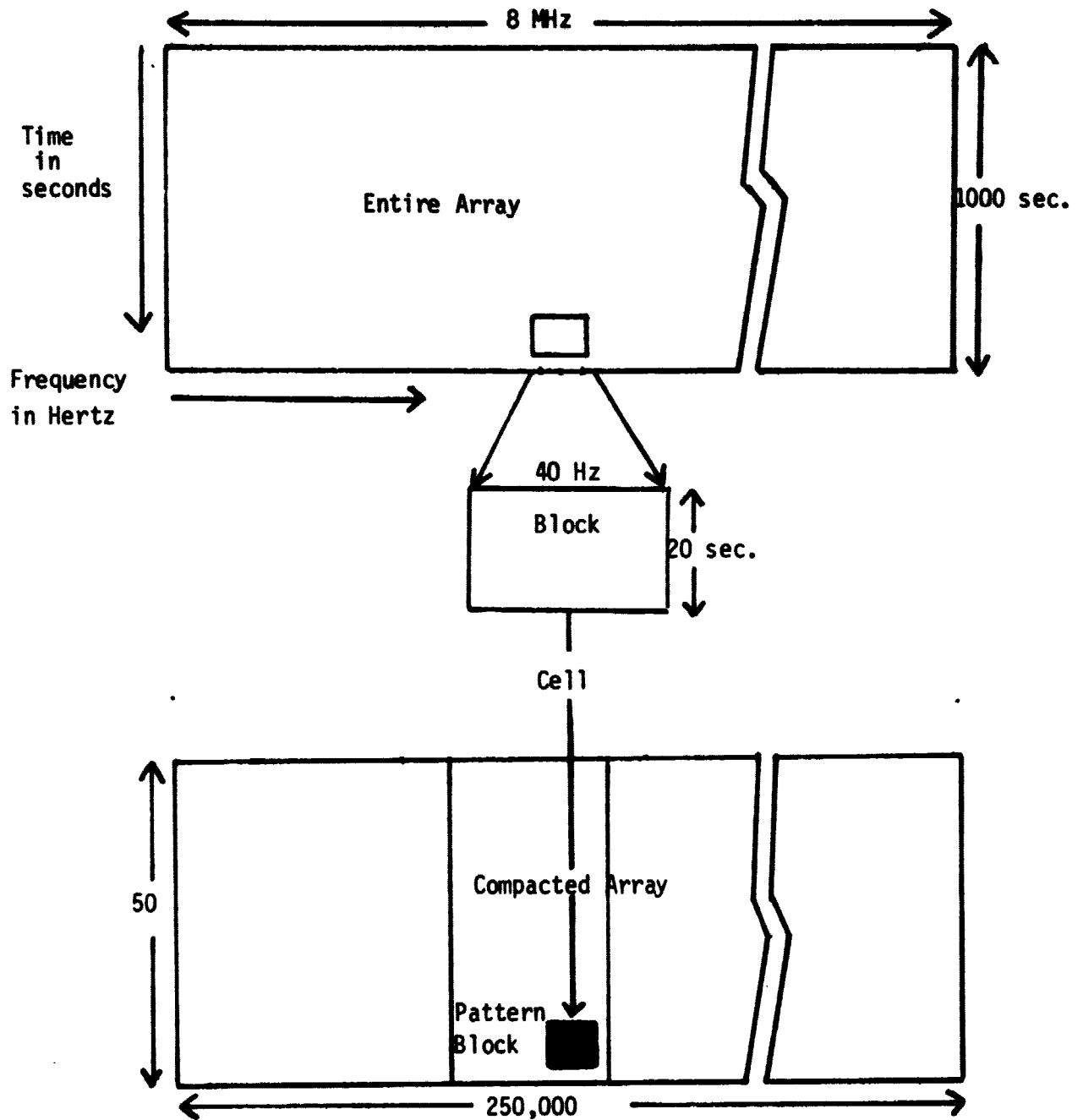


Figure 3.12 - Definitions of the Terms used for Cluster Analysis



clusters such that the elements within a cluster have a high degree of "natural" association. The two ways we define this natural association in our application is continuity in time and/or frequency of similar pattern points.

We view a pattern point as a point in the compacted data array which has associated with it the time and frequency coordinates of the cell it corresponds to. Each bit of a pattern point is set by one of the tests described in stage I and has value either zero or one. If a signal is present which sets some particular bits for some length of time and/or across some frequency width, then the pattern points which have these bits set will lie along meandering lines or be grouped in other types of clusters in the compacted frequency-time domain. Hence, we need algorithms which will discover various spatial relationships between similar pattern points. In order to obtain this objective, we use a two step cluster algorithm which consists of a "smart" histogram and a modified single linkage cluster algorithm.

5. Overview - The method by which a pattern block is examined for the presence of a cluster is straightforward and will be described briefly in this section, with more detail being given in Appendix A3.11.

We recall that a cluster is defined as a close spatial association of a number of pattern points belonging to the same pattern class. As an example of a cluster consider Figure 3.13 which shows a portion of one pattern block indicating all pattern points having bits 1 through 5 ; i.e., members of class (1, 1, 1, 1, 1, \*, \*, \*, \*). Since there is a probability of  $1/2$  that any of these bits will be on due to the fluctuation of the noise background, the probability of them all being on in the same pattern point is  $1/32$  due

ORIGINAL PAGE IS  
OF POOR QUALITY

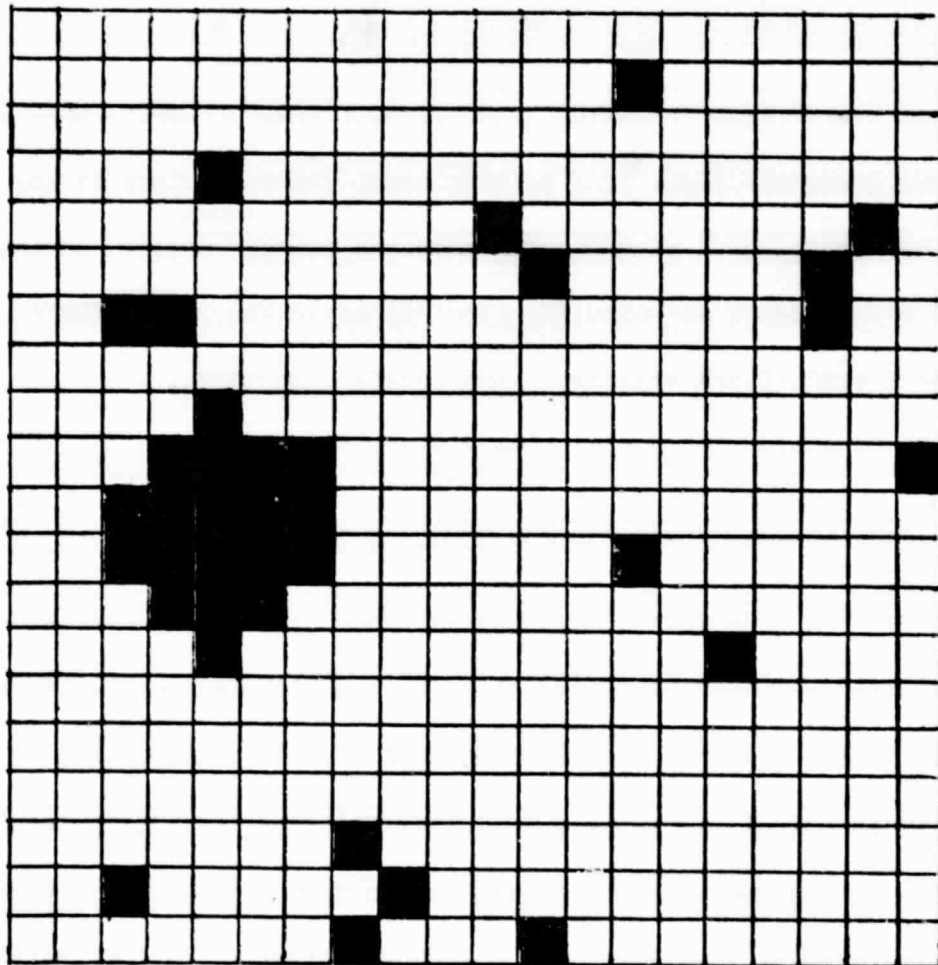


Figure 3.13 - A Cleaned Pattern Block

to the background alone. The algorithm searches for deviations from the statistical noise field by specifically searching for a group of N "on cells" each being within a spatial distance  $d$  of at least one other member of the cluster. The algorithm detects such a pattern in two stages.

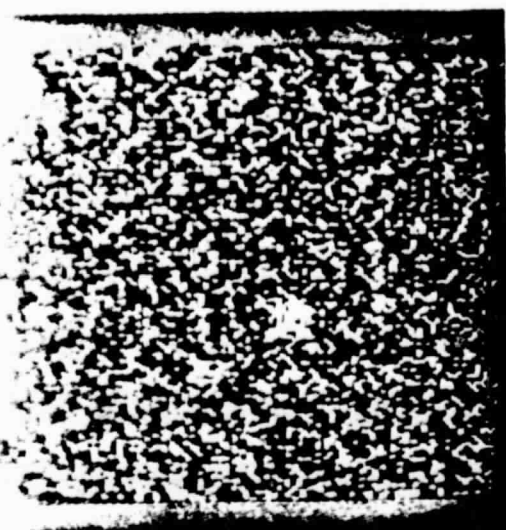
The first stage is called the histogram stage. Here a dedicated hardware processor looks at a pattern block and determines if any pattern class has an "excess" of members within the block. Classes having a large excess will then be selected for stage II analysis. Furthermore, an extremely high excess will alert the central computer.

In the second stage analysis, a dedicated software processor first "cleans" the pattern block of all pattern points not in the class of interest. The result of this cleaning procedure is a 2-dimensional binary representation as shown in Figure 3.13. The binary representation, whose constituents are also called "cells", is then combed through and a list is kept in the software processors memory of all closest pairs of "on" cells. These pairs are deemed to be the likeliest centers of clustering, and are used as seeds around which clusters may be recognized. The algorithm that is then performed is a variant of the single linkage algorithm (Andenburg, 1973, and Ton and Gonzales, 1974), which collects adjacent points, that is, points within a predetermined distance  $d$ , into clusters. If the processor finds itself able to extend the cluster spatially in a particular direction or in all directions such that a highly significant number of points are amassed, then a cluster has been discovered. At this point, the central computer is alerted and the data pertinent to the cluster, which includes both the raw

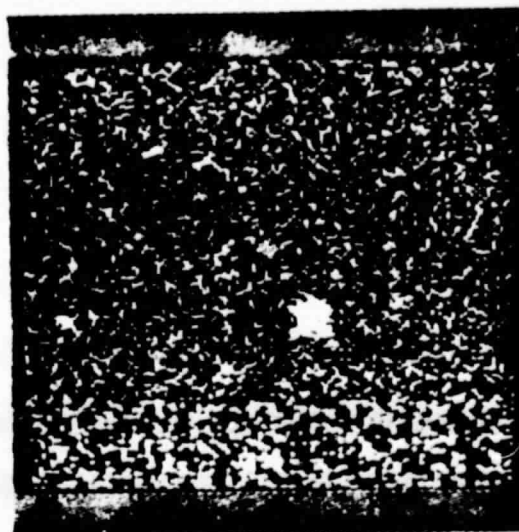
data and the cell representation, is archived and available for display. Simulations were conducted with high-level languages software that were very effective. These are shown in Figures 3.14 - 3.16.

The great advantage of this algorithm is that it is able to search for structure across any entire pattern block, which is effectively a 6.4 kHz by 1000 second record. The compactification scheme has bought us the real-time ability to do such extensive spatial searches. This capability is far in excess of what is possible through the use of dedicated processing of an equivalent 8 billion pixel image. In addition the dimensions of phase, polarization, and several types of coherence (temporal, spatial, phase) are included in the pattern point compaction scheme.

ORIGINAL PAGE  
BLACK AND WHITE PHOTOGRAPH

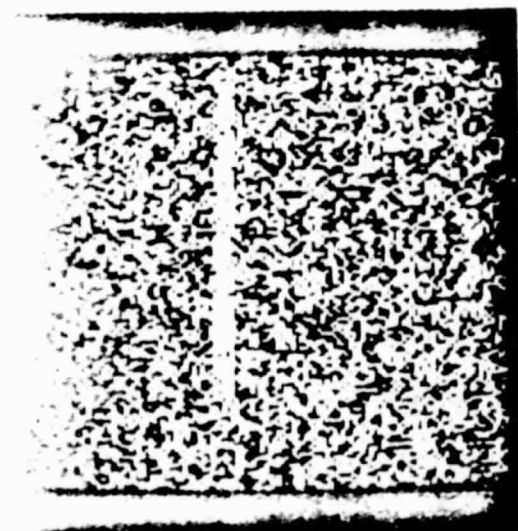


(a) Before

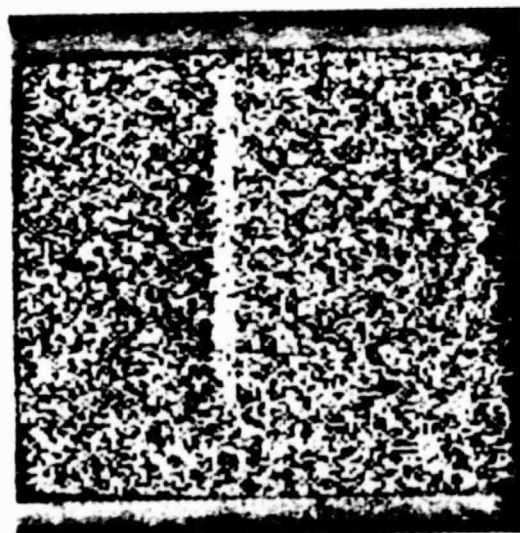


(b) After

Figure 3.14 — A cluster is detected and displayed by highlighting cells that have 12 or more neighbors within a distance of 2 units.

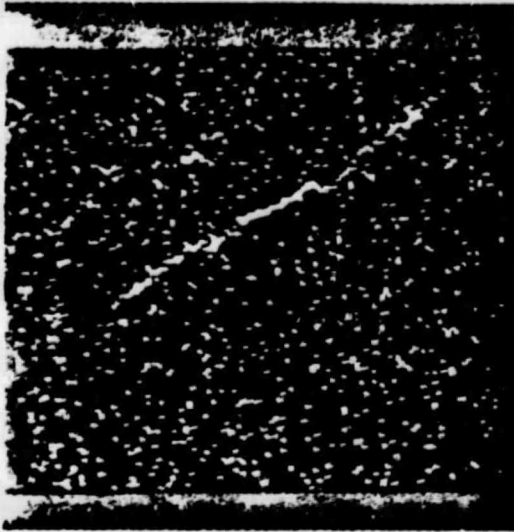


(a) Before

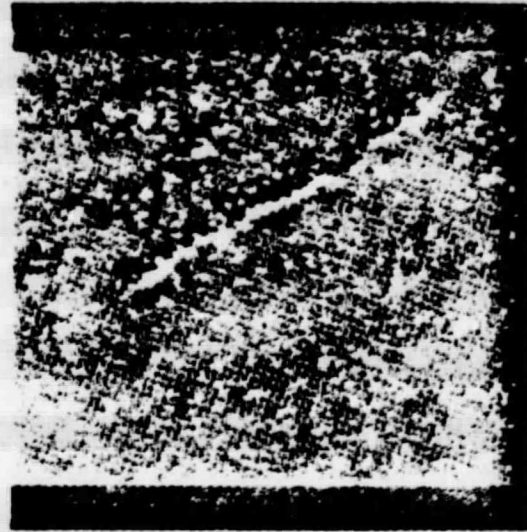


(b) After

Figure 3.15 — In this densely populated pattern class (1,\*,\*,\*,\*,\*,\*,\*) the vertical feature was discovered by testing for cells that are completely surrounded, i.e., that have 8 neighbors within a distance of 1 unit.

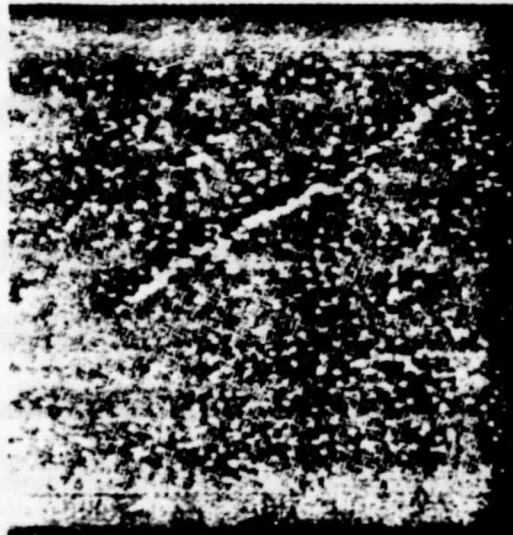


(a) Before



(b) After First Pass

Figure 3.16 — To isolate a tenuous feature, we first require that a cell have at least 2 adjacent neighbors to survive highlighting (b), then in a second pass (c) we require that the survivors have three neighboring survivors within a distance of 2 units.



(c) After Second Pass

8  
Abramowitz, M., and Stegun, I.A., editors, Handbook of Mathematical Functions, National Bureau of Standards Applied Mathematics Series 55, 1964.

Alexandridis, N.A., "Walsh-Hadamard Transformations in Image Processing" AFOSR Report No. 71-1362, March 1971.

Andenburg, Michael R., Cluster Analysis for Applications, Academic Press, 1973.

Bendat, J.S., Principles and Applications of Random Noise Theory, John Wiley, New York, 1958.

Beran, M.J., and Parrant, M.J., Theory of Partial Coherence, Prentice-Hall, 1964.

Bergland, G.D., "A Guided Tour of the FFT," Digital Signal Processing, (ed. by Rabiner, Rader), IEEE, 1972.

Blachman, N.M., "Sinusoids versus Walsh Functions," Proceedings of the IEEE, Vol. 62, NO. 3, March 1974.

Born, M. and Wolf, E., Principles of Optics, Pergamon Press, 1970.

Bracewell, R., The Fourier Transform and Its Applications, 2nd Edition, McGraw-Hill, New York, 1978.

Chen, Kok, Ph.D. dissertation in progress, Department of Electrical Engineering, Stanford University, 1980.

Cochran, W.G. and Cox, G., Experimental Design, John Wiley and Sons, Inc., New York, 1962.

Cohen, M.H., Radio Astronomy Polarization Measurements, Proc. IRE, Vol. 46, p. 172, January 1958.

Deans, S.R., The Hough Transform from the Radon Transform, preprint August, 1979.

Deans, S.R., preprint, August 1979.

Dixon, R., "A Search Strategy for Finding Extraterrestrial Beacons", Icarus, 20, 187-199, 1973.

Duda, R.O. and Hart, P.E., Commun. Ass. Comput. Mach. 15, 11-15, 1972.

Franck, J.B., Hough, P.V.C., and Powell, B.W., Nuclear Instruments and Methods 20, 387-392, North-Holland, 1963.

Fu, King S., Digital Pattern Recognition, Springer Verlag, 1976.

Gel'rand, I.M., Graev, M.I., Vilenkin, N.Y., Generalized Functions, Academic Press, New York, 1966, Vol. 5.

Handbook of Tables for Probability and Statistics: 2nd Edition, pp. 247-255, Chemical Rubber Co., Cleveland, Ohio, 1968.

Hanson, K.M., .A-UR 78-1597, Medical Physics (to be published, 1979)

Harmuth, H.F., Sequency Theory, Academic Press, 1977.

Hogg, R.V. and Craig, A.T., Introduction to Mathematical Statistics, 3rd Edition, McMillion Publishing Co., Inc., 1970.

Hough, P.V.C., and Powell, B.W., A Method for Faster Analysis of Bubble Chamber Photographs, Proceedings of an International Conference on Instrumentation, Berkeley, Sept. 12-14, 1960, Interscience, N.Y., 1961, pp 242-245.

Hough, P.V.C., Method and Means for Recognizing Complex Patterns, U.S. Patent 3069 654, 1962.

Ko, H.C., On the Reception of Quasi-Monochromatic, Partially Polarized Radio Waves, Proc. IRE, vol.50, p.1950, Sept. 1962.

Kraus, J.D., Radio Astronomy, McGraw-Hill, New York, 1966.

Lee, Y.W., Statistical Theory of Communication, John Wiley, New York, 1960.

Machol, Robert E., Two Systems Analyses of SETI, unpublished apaper on file in SETI Program Office, NASA-Ames, Moffett Field, Ca., 19 .

Middleton, David, An Introduction to Statistical Communication Theory, McGraw-Hill, New York, 1960.

Middleton, David, Topics in Communication Theory, McGraw-Hill, New York, 1965.

Oliver, B. and Billingham, J., Project Cyclops, NASA-Ames, Moffett Field, CA.

Otnes, R.K. and Enochson, L., Applied Time Series Analysis, Volume 1, John Wiley, New York, 1978.

Radon, J., Ber. Verh. Sachs., Akad. 69, 262-277, 1917.

Scheffe, H., The Analysis of Variance, John Wiley and Sons, Inc., New York, 1959.

Schreiber, H. and Sandy, G.F. (ed.), Applications of Walsh Functions and Sequency Theory, IEEE, 1974.

Schwartz, M., Bennet, M., and Stein, M., Information Transmission Modulation and Noise, McGraw Hill, New York, 1966.

Seeger, C., "Notes on Search Space", In The Search for Extraterrestrial Intelligence, Philip Morrison, John Billingham and John Wolfe, eds., Superintendent of Documents, Washington.

Shapiro, S.D., Computer Graphics and Image Processing 8, 219-236, 1978.



Shapiro, S.D., Pattern Recognition 10, 129-143, 1978.

Shapiro, S.D., and Iannino, A., IEE Transactions Pattern Anal. and Mach. Intelligence, PAMI-1, 310-317, 1979.

Sklansky, J., IEEE Transactions Comput. c-27, 923-926, 1978.

Sogliero, G., private communication, 1980.

Ton, Julius T., and Gonzales, Rafael C., Pattern Recognition Principles, Addison-Wesley Publishing Company, Inc., 1974.

Whalen, A.D., Detection of Signals in Noise, Academic Press, New York, 1971.

## Chapter 4

### Oasis System Architecture

The centipede was happy quite,  
Until the Toad in fun  
Said, "Pray which leg goes after which?"  
And worked her mind to such a pitch,  
She lay distracted in the ditch  
Considering how to run.  
--Mrs. Edmund Craster

## 4.1 - Introduction

### Purpose of the System

The purpose of the Oasis Signal Detector is to recognize signals in noise and to identify them in so far as is possible. In addition, the system:

- Archives interesting data.
- Allows the operator to become involved in decision-making and to introduce signal identification tests of his own.
- Provides the operator with visual and audio feedback.
- Oversees the operation of the entire receiving system, including the sequence of events, such as changing to a new target.

The system has been described in the previous chapter as a set of algorithms (Pulse Detector, Carrier Wave Detector, Numerical Battery of Independent Tests, Cluster Detector). These algorithms comprise the Real-time Processing System; there are, in addition, three other systems in the detector: the Total Observation Archive, the Operator Interface, and the Radio Astronomy System, as shown in Figure 4.1. Now we will describe the digital hardware implementation of these systems. First there will be an overview of the architecture and data flow within the entire detector, followed by a more detailed examination of the processors, with schematics presented wherever necessary. The chapter ends with a summary of the speed, capacity, and cost of the planned equipment.

The system employs both conventional digital computers and special purpose digital processors to implement the algorithms. The reasons for the

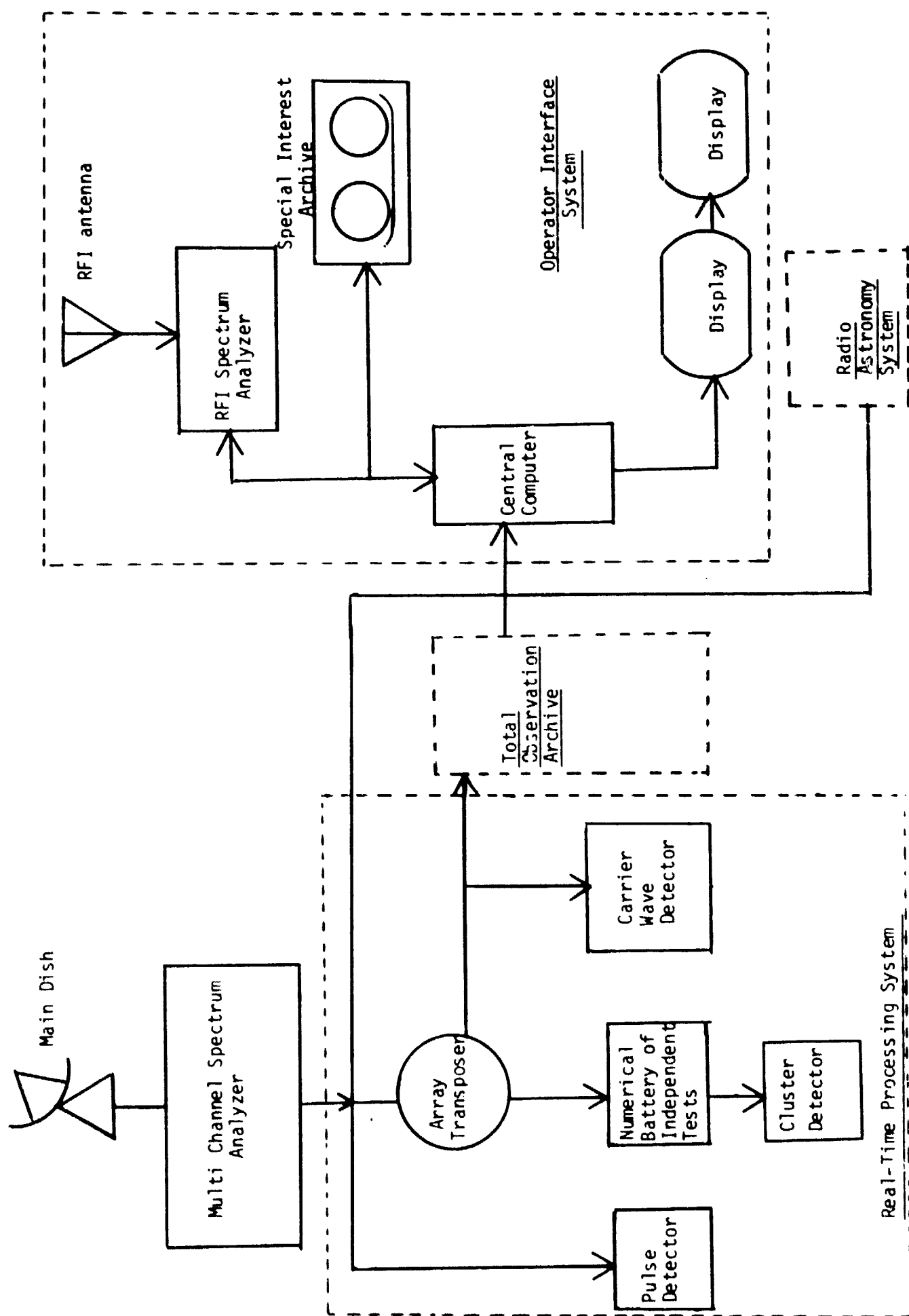


Figure 4.1 - The Multichannel Spectrum Analyzer and the Four Systems That Comprise the Oasis Signal Detector

particular choice of technology is discussed in detail in Appendix A4.1. While the selection of algorithms was somewhat constrained by the availability of compatible technology, the hardware feasibility study which accompanied this design project had its largest impact in placing limitations on the sizes of the random access memories available to each processor.

### Architectural Strategy

The strategy and philosophy underlying the hardware design which follows can be stated rather simply. If proven off-the-shelf components and processors exist to do a task, we use them. Otherwise, as in implementing several of the computational algorithms, we have found it reasonably straightforward to design custom-made processors, all of which may be constructed simply using current or near-current technology.

When the data rates and computation rates exceed that which is possible with conventional integrated circuit technology, we use multiple processing, and/or parallel processing, rather than resorting to higher speed processing (e.g., greater than 10 MHz), and thus avoid immature and somewhat incompatible integrated circuit families.

Fortunately, the very nature of the MCSA's output scheme (8 megawords/sec) allows all special purpose processors located downstream of this data flow to operate with a fundamental clock frequency of 8 MHz or 10 MHz,<sup>1</sup> which is

---

1 The reason for 10 MHz is due to the redundant processing of overlapped data as discussed in the discussion of Figure 3.9, detail 1.

within the range of TTL technology. Still, there are cases where the number of computations to be performed at this rate is very large, and in these cases parallel processing is employed, either with special purpose hard-wired devices (e.g., in the CWD) or else using a set of microprocessors (e.g., as in the CD). All processors as well as all the major subunits of the Oasis Signal Detector make extensive use of the techniques of pipelining and double buffering ("ping-ponging"). Each stage of the pipeline introduces a delay in the detection process, but each stage is allowed to operate at 8 MHz or 10 MHz. In some stages of the pipeline the data is reduced or "compacted" as well. By this means, in the second stage of the Realtime Processing System, the data rates become much more manageable. Here, data compaction by a factor of over one thousand (via the NBIT routine) reduces data rates to 12.5 kHz and smaller words, allowing a sensitive, non-deterministic search for patterns to take place by way of the Cluster Detecting firmware routines of a microprocessor.

#### Large Memories and "Second Looks" at the Data

An important question, and one studied in depth by the group, was the feasibility of maintaining a large online memory containing all of the raw data from a target (.1 to .5 terabit). We determined that to store such a data base was possible, yet such a data base could not be designed to afford convenient random access to the data for preliminary processing. While it would have been desirable from both an architectural and algorithmic point of view to have a large random access data base through which all processors could comb, backtrack, and focus in or out of, such an entity was not found feasible. Magnetic disc technology has not reached the storage density or capacity,

magnetic tape has not reached the required access rate, and erasable digital optical storage is yet immature and has not solved error rate problems at high densities, nor are any of these constraints expected to be sufficiently lifted within the next five years.

We did, however, design a large memory in the form of a "Super High Bit Recorder Helical Tape Transport", which serves, not as a primary data base, but as an archive, in both a short term and long term sense. The data taken from every target is stored temporarily in this "entire observation archive" while the dedicated processors (PD, CWD, NBIT) concurrently run deterministic tests on an identical stream of data. The tests are considered deterministic in the sense that these tests simply act and report their findings over and over again; never altering their behavior in response to their own results. After this deterministic search has transpired for a particular target and any outstanding results of the tests have been gathered, a "second look" is taken at the data, using the results of the PD, CWD and NBIT processors as a guide to examining the archive. This is shown in Figure 4.2.

The concept of taking a close second look at promising portions of the data after an initial evaluation has been made is a natural and convenient solution to the large memory problem state above. In this way, all of the data is made available to powerful, sensitive, and non-deterministic tests, such as are provided by the clustering algorithm and the operator's own bank of knowledge, skills and tools. Nor is the concept of a second look new. It has often been suggested by those working in the field,<sup>2</sup> and it emerged independently as well from the ideas of this design team.

The data compaction of .1 terabits to .1 gigabits allows the system to automatically take a second look at the data. The second look comes in two

<sup>2</sup>Seeger, Stull, 1979, personal communication with members of the summer study.

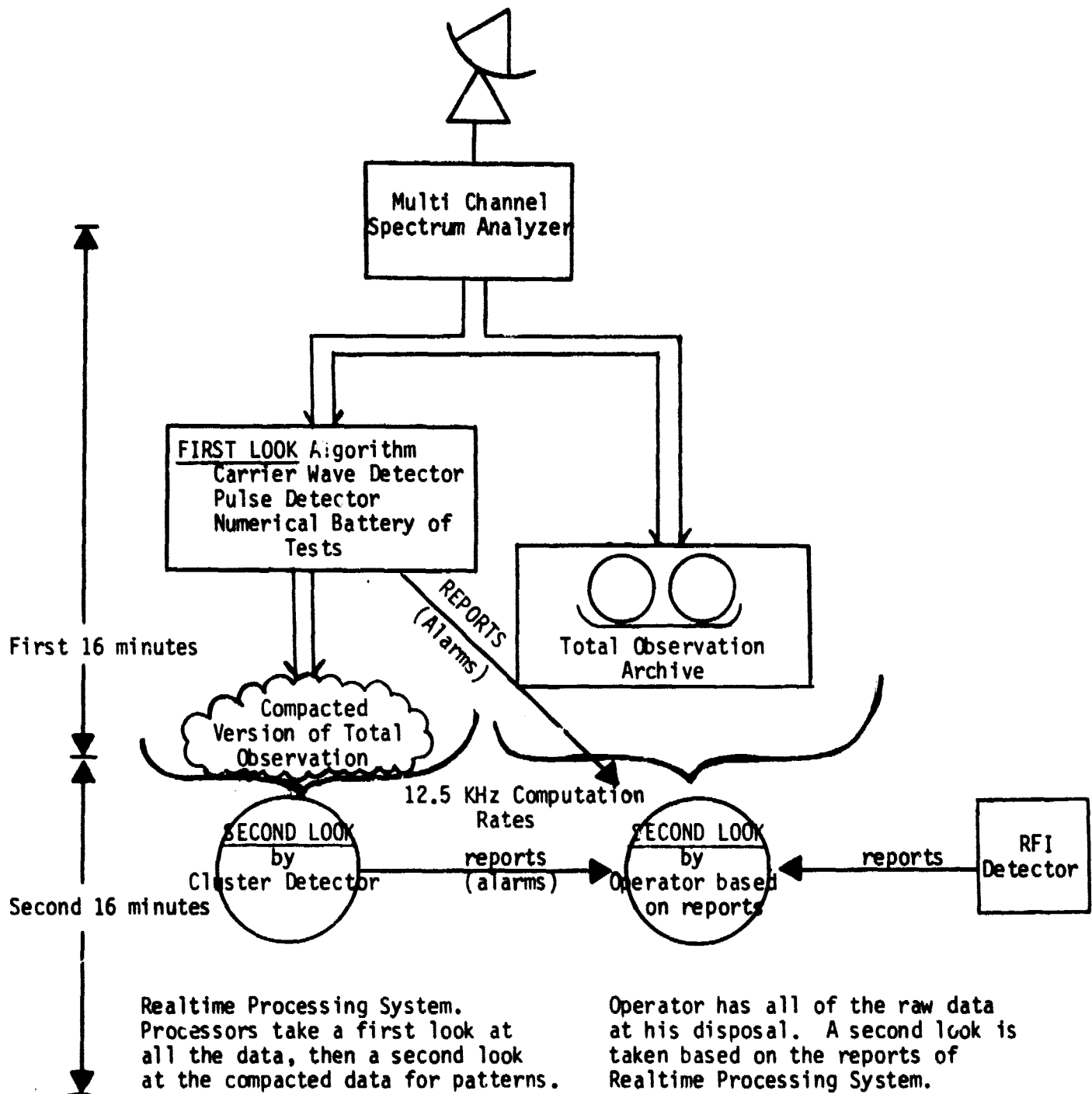


Figure 4.2 - Use of the "Second Look" strategy, with both machine and operator getting a second look at the data. The operator is used as the ultimate pattern recognizer.



forms: the histogramming and the adjacency search algorithms of the Cluster Detector.

The histogramming tests each 9 bit "pattern point" coming from the NBIT detector at a 12.5 KHz rate for membership in any of 512 "pattern classes". It can be seen that by compressing the data here we have gained the ability to subject it to many more tests.

The operator is also given a second look at the data. As will be discussed in the next chapter, it is felt that the operator can and should play an integral role in signal detection since he possesses the most advanced and sophisticated algorithms for pattern recognition possible. The previewing of data by the hardware processors rather than the operator is necessitated by the fact that the raw data, if presented on a 22-inch-tall standard display, would require a display 43 miles in length.<sup>3</sup>

Instead, the operator is able to concentrate on alarms from, and possible patterns detected by the real-time processing system. To aid him in his task, he is provided with exclusive use of three archives, an advanced image processing unit, and a bank of software data reduction routines, as well as aid from a radio frequency interference detector and a knowledgeable radio astronomer. At the end of the 16-minute period, as the system moves on to observe the next star, a decision is made, either by the operator, or by the central computer, to save or erase the prior observation. The operator's role as a signal detector is shown in Figure 4.2. Comments on the criteria for the preservation for an observation will be given in the next chapter.

---

<sup>3</sup> Assumes each 22"x22" area displays 512x512 pixels of 16 levels, and one observation consists of  $8 \times 10^6$  Hz x 1000 seconds x 2 polarization x (1 real + 1 imaginary part) x 16 levels.

### Advantages of the Architecture Chosen

The system is highly modular, with little or no communication taking place between processors sharing the same stage in the pipeline. Input to processors arrives at a fixed rate on a common bus, and output is typically limited to the signaling of an alarm through a priority interrupt to the central computer. The computer acquires identifying information regarding such a processor's alarm by reading the output register of the processor without establishing any two-way communication.

There is great advantage in keeping different functions physically separate and communications between modules minimal. In addition to providing a substantial increase in computation power through parallel processing, this scheme allows for easy modification of the system, through addition or removal of processors. Processors may be easily monitored for failure, and the failure of one will not bring down the entire system. The central computer simply removes the processor from the priority interrupt system, and, if necessary, terminates the pipeline at this processor.

Such modularity provides for a great deal of options in the implementation of the Signal Detector. In Figure 4.3 we show the basic pipeline, each level of which is a set of signal detectors. One may start at the bottom, throwing away modules without affecting the performance of the modules on the left, right, or above. At each turn, one is left with a functioning signal detector, albeit one with reduced sensitivity to some class of signals.

It should be noted that the most fundamental algorithms are located in the first two lines of this figure, namely the Pulse, Carrier Wave, and the nine NBIT algorithms. Thus, a minimal system configuration would still contain the

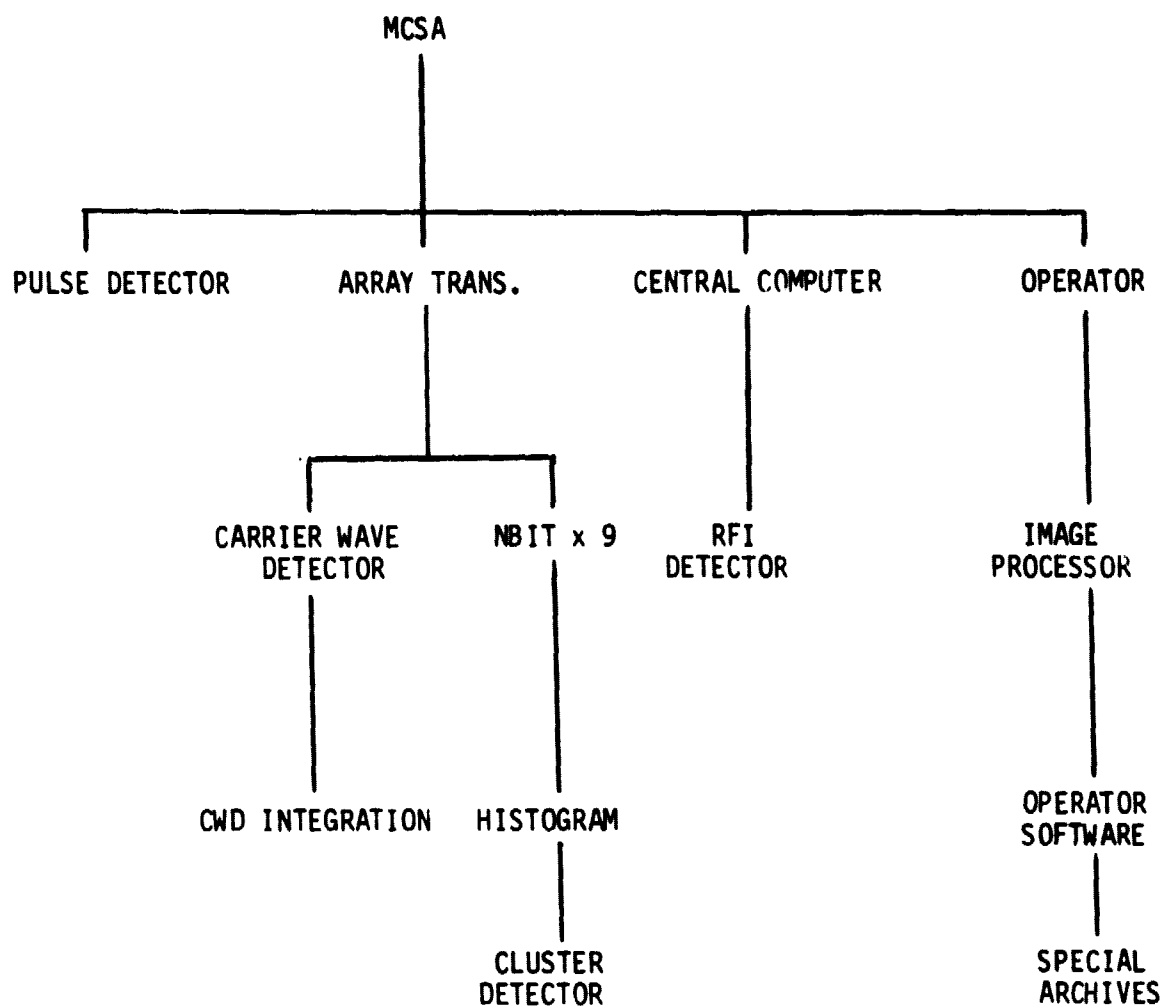


Figure 4.3 - A Delineation of the Pipeline of the Oasis Signal Detector. Modules depend only on those directly above. Each module excluding Array Transposer is a signal detector.

most sensitive, and important algorithms. Further, any of the particular NBIT algorithms may be replaced by any alternative bit-setting algorithm, with the entire system above and below this stage of the pipeline remaining unaffected.

An advantage of the system architecture of a more primary nature is the fact that all components to be described in this chapter are proven, off-the-shelf items. Though some of the hardware is state-of-the-art within its class, namely the high density discs, helical tape transport, and display processor, as well as the integrated circuits for particular integer multiplies, the classes themselves are mature in the sense of standardization, reliability, economics, and the very fact that corporations supporting such products are well-established, and reasonably long-lived. Appendices A4.1 and A4.2 deal further with the choice of technology, and alternative ways to implement the OSD modules based, for example, on smaller or larger MCSAs, and the use of fewer modules.

#### 4.2 - The Real-Time Processing System

##### Input to the Oasis Signal Detector

The MCSA's bottom stage, consisting of  $140 \times 56 = 7840$  BPF units, each with associated RAM memories, provides the input data to the signal detector. As can be seen in Figure 4.4, this stage produces, at each RAM memory, 1024 output words, each representing a channel of width of approximately 1 Hz. With 7840 BPFs of 1024 "channels" each, there are, in total, 8 million outputs. These outputs are produced each second, so that an 8 MHz bandpass input to the MCSA may be totally analyzed to 1 Hz resolution and bussed to the OSD in one second.

Input: 8 MHz Baseband Signal  
 Sampling: 10 MHz Quadrature Oversampling

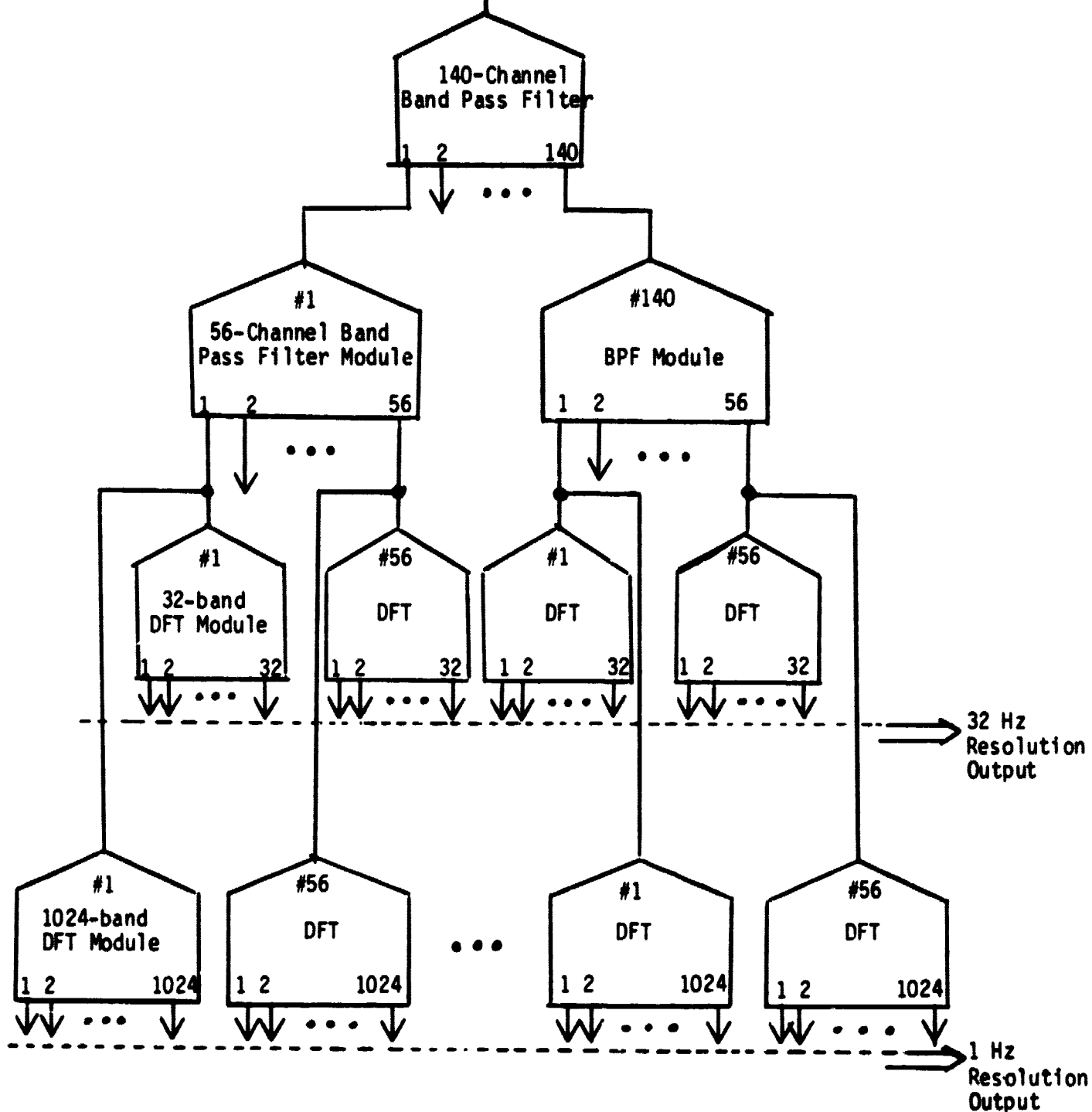


Figure 4.4 - MCSA Architectural Block Diagram, with each Band Pass Filter Employing Prime Factor FFTs, DFTs, and Finite Impulse Response Filters.

During one of these time intervals, all the RAMs fill in parallel. The results of the Prime Factor Algorithm (Narasimha, Peterson, and Narayan, 1978)<sup>4</sup> leaves the data skewed across the RAM in a fashion similar to the results of an FFT operation (Bergland, 1969). By providing two RAMs at such stages, one RAM may be written into in a skewed fashion, while from the other, data is read out from the previous time interval in a sorted manner. Upon leaving the final DFTs, the data enters the Signal Detector. Expanded block diagrams of the OSD showing all major components with associated data rates and memory capacities are shown in Figures 4.5 and 4.6.

The data is read out at 8 Mwords per second, with each word consisting of 16 bits real and 16 bits imaginary, and is clipped after readout to 4 bits real and 4 bits imaginary, then multiplexed (box labeled MUX in Figure 4.5) with the output of the second identical MCSA operating at an orthogonal polarization. This yields a composite 16-bit word consisting of 4 real bits (polz. X), 4 bits imaginary (polz. X), 4 bits real (polz. Y), 4 bits imaginary (polz. Y). As discussed in Chapter 3, each quantization level of the new 4-bit values represents .3σ of the average signal amplitude. The clipping is accomplished with an 8-bit to 4-bit ROM look up table.

### Array Transposer

The output of the multiplexer, the 8 Mwords/sec data flow branches in two directions. A scan (a one-second observation from the MCSA, 8 million 16-bit composite numbers) read out of the multiplexer is bussed word by word to the pulse detector, which performs its computations immediately and directly on each scan. The scans are also written into a device known as the array transposer.

<sup>4</sup>See references Chapter 3

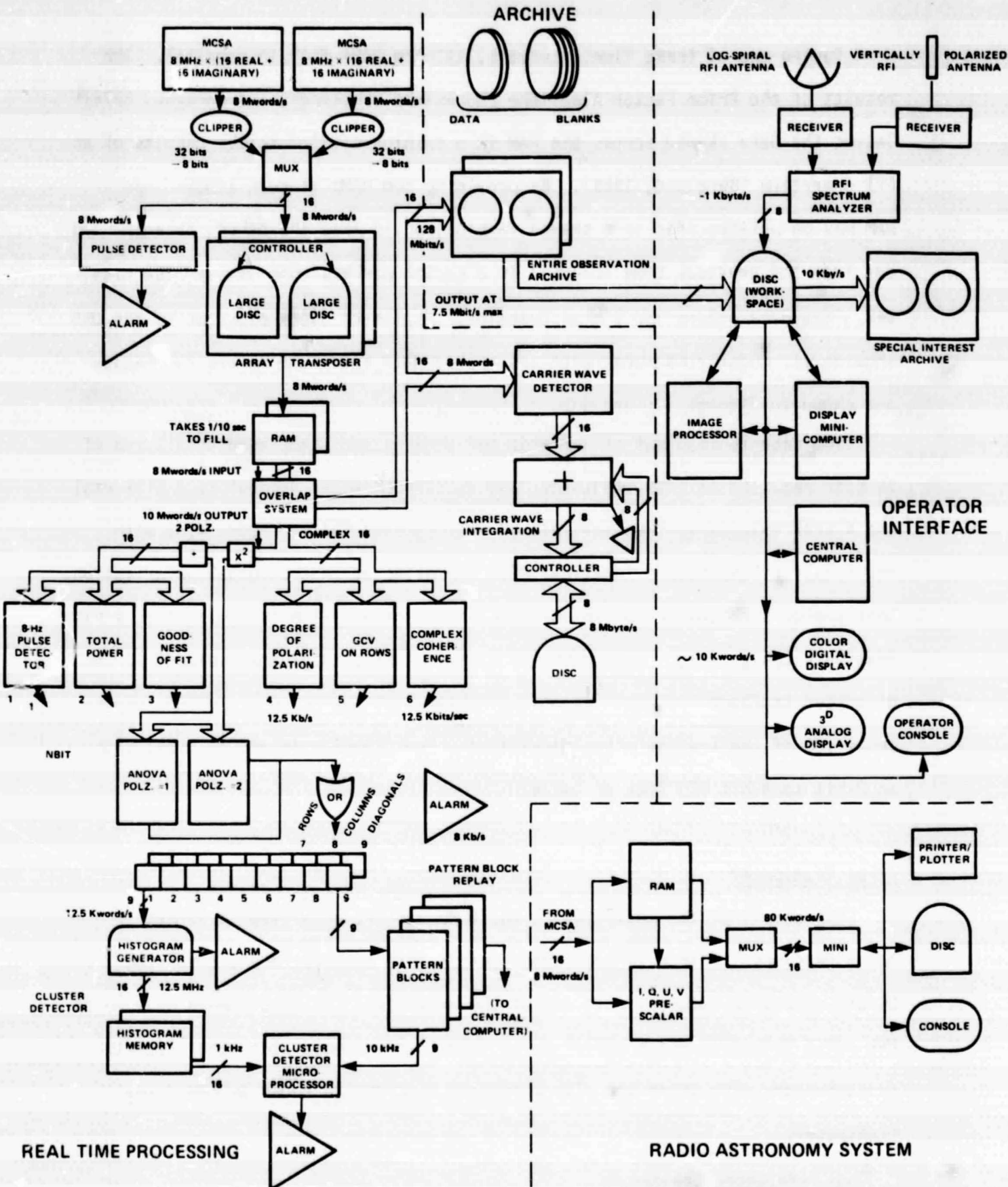


Figure 4.5 Oasis Signal Detector Block Diagram With  
Rate of Data Flow Between Components

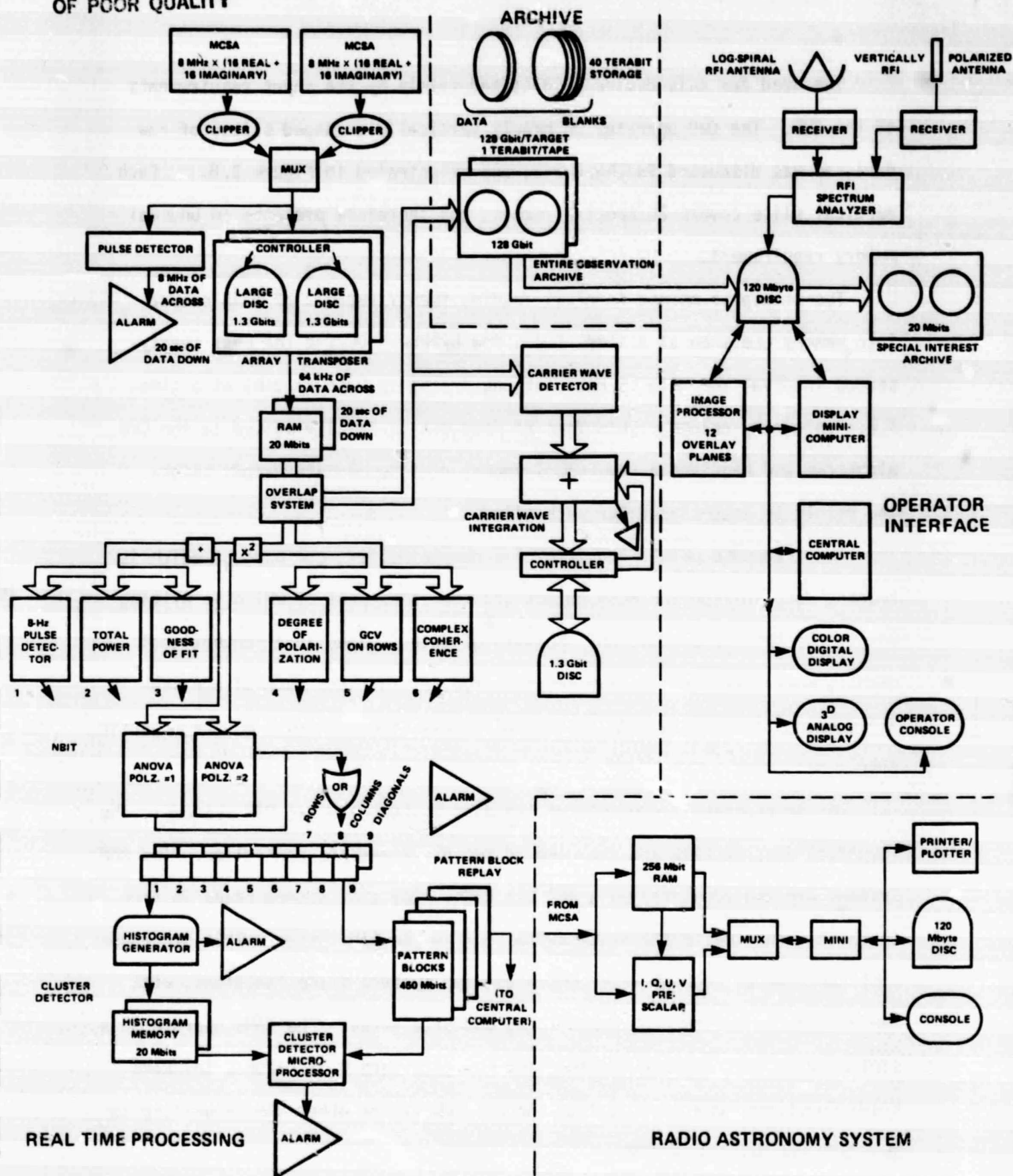


Figure 4.6 Oasis Signal Detector Block Diagram With Storage Capacities of Memory Components



The need for this device is dictated solely by the input requirements of the CWD. The CWD operates on nearly vertically arranged slices of raw data, as was discussed in Chapter 3, and illustrated in Figure 3.8. Each vertical slice covers 20 spectral scans, and therefore presents an unusual memory requirement.

The transpose method involves reading twenty seconds of MCSA output into memory one scan at a time, i.e., row by row. During the next twenty-second interval the data is read out one frequency (one column) at a time, i.e., column by column. This form of read out is ideally suited to the CWD algorithm and requires a very small amount of intermediate memory before the result is accumulated for each slice.

The transpose method consists of a double buffer, one part of which is reading data in from the MCSA output and the other part of which is writing out to the CWD implementation. Read-in of data is done in a conventional manner, where one scan is read in at a time, so that the data of one scan is arranged contiguously around one set of tracks (one cylinder) on the magnetic disc.

The transposer is composed of two disc systems and two RAM systems. The identical disc systems are composed of two discs and one controller. The RAM systems are two identical  $20 \times 10^6$  bit RAMs. One disc system reads in from the MCSA output while the other writes out to the RAM system. This procedure continues for 20 seconds after which the two systems trade functions, with the disc which was reading now writing and vice versa. The RAMs operate on a similar basis but on much shorter time scales. Each 0.13 seconds the RAMs trade functions, one reading from the disc while the other RAM is writing to the CWD and NBIT algorithms.

This double-buffering of systems allows for the transposing of twenty scans from row-by-row access to column-by-column access in the following manner. Data is written from one of the disc systems to one of the RAMs in partial rows. These are  $64 \times 10^3$  Hz strips of an MCSA output scan ( $1/130^{\text{th}}$  of a scan). Within the disc system that is writing out, there exist two disc drives and a controller. One disc reads out a partial scan to the RAM while the other disc seeks out another partial scan, positioning its read-head appropriately. The amount of time required to perform this "seek" is equal to the amount of time required to write out a partial scan. It is in this way that time is bought to fill one of the comparatively small RAMs with a  $64 \times 10^3$  Hz x 20 second piece of the total 20 seconds of data. Once the data has filled the RAM, the RAM switches to the write mode and the data is written out, twenty-element column by twenty-element column. This constitutes the transposed data flow.

In the read and write modes described above, all data transfer rates are  $8 \times 10^6$  words/sec., where a word consists of 16 bits, 4 bits real and 4 bits imaginary for each polarization. It should be noted that the data rate out of the MCSA is maintained at all times. The net effect of the disc-double-buffering is to delay the "real-time" arrival of the data by 20 seconds as well as to transpose it.

#### Advantages of the Transposition Scheme

The transposition scheme is necessitated by the requirements of the CWD algorithm. There are two additional and very important benefits of the transposing scheme. The first is the block-by-block transmission of the raw data to the NBIT processors. With this transmission scheme, the 800 raw data words

that are used to produce one bit of the 9-bit number which represents a block all arrive sequentially. Hence, none of the partial results of an NBIT test on a particular block need be saved after that block is processed.

It should be recognized that there is an alternative processing scheme for the NBIT algorithms that is eliminated by using the array transposer. This scheme is called the "on-the-fly" implementation. It consists of taking the raw data from the MCSA, one scan at a time, and computing the results for all pattern blocks in parallel. This on-the-fly scheme would require accumulators for each block during the processing of twenty scans. There are  $2.5 \times 10^5$  blocks being parallel<sup>5</sup> processed and the amount of accumulator memory required would be approximately  $3 \times 10^8$  bits. This is not an unreasonable amount of memory, but nonetheless the need for its presence is completely eliminated by the transpose system.

#### Pulse Detector Implementation

The algorithm implemented here, as described in Chapter 3, is an examination of the incoming scans under various resolutions ranging from 1 Hz, 2 Hz, 4 Hz, ... ,  $2^{23}$  Hz = 8 MHz. This is accomplished in the following manner. First the incoming signal is converted to a total power value by summing the squares of 4 complex components. The components used in this operation are "lookup" ROMs for squaring and MSI integer adders. Then 24 such adders integrate the data into 24 accumulators. At the end of  $2^{i-1}$  integrations, the  $i^{\text{th}}$  accumulator is compared against a threshold. If there

---

5 "Parallel" in the sense that the processor would work on each block only 1/20 of the total amount required to complete that block, and then move on to the next block. Only after 20 visits to each block is this parallel computation completed.

is a significant excess of total power, the threshold is exceeded, and an alarm is sent to the central computer.

This hardware scheme is shown in Figure 4.7. The clock scheme for the adders is simply an 8 MHz oscillator synchroized with the MCSA readout, and this clock drives a 24-bit counter. Each bit of the counter is an enable signal (e.g., rising edge due to  $0 \rightarrow 1$  transition) to the thresholding function of a unit. All adders add upon each cycle of the clock, and, at each  $2^{i-1}$  cycle, the  $i^{\text{th}}$  bit of the counter experiences a rising edge, and allows a threshold and accumulator clear to take place. The expected total power from  $2^i$  summed channels is a constant due to the clipping procedure, and thus the threshold values are likewise constants, hardwired into digital comparison chips which check for "greater than" situations.

When such a situation exists in any of the 24 units, an or-ing procedure allows a single alarm (one-shot interrupt pulse) to be sent to the central computer's priority interrupt system. The unit experiencing the excess integrated power latches its own identifying number "i" into a register connected to the computer's IO interface. An additional pair of registers, the 24-bit counter already discussed, and a scan number counter are also similarly connected to the computer, and these three registers serve to uniquely describe the location of the anomaly in the raw data, for the operator interface system.

The detector, as described above, performs the coarse resolution pulse search on one scan at a time. The detector includes, as well, a second, identical unit that performs this test on an integrated scan (see Chapter 3). This integrated scan is accumulated at an 8 MHz rate into a large RAM memory  $16 \text{ bits} \times 8 \times 10^6 \text{ bits}$ .<sup>6</sup> At the end of 1000 seconds (1000 scans) this

---

<sup>6</sup> While RAMs of this size are not manufactured per se, in a serial access utilization such as above, there is little difficulty in stringing together many 256K RAMs, without encountering additional addressing problems.

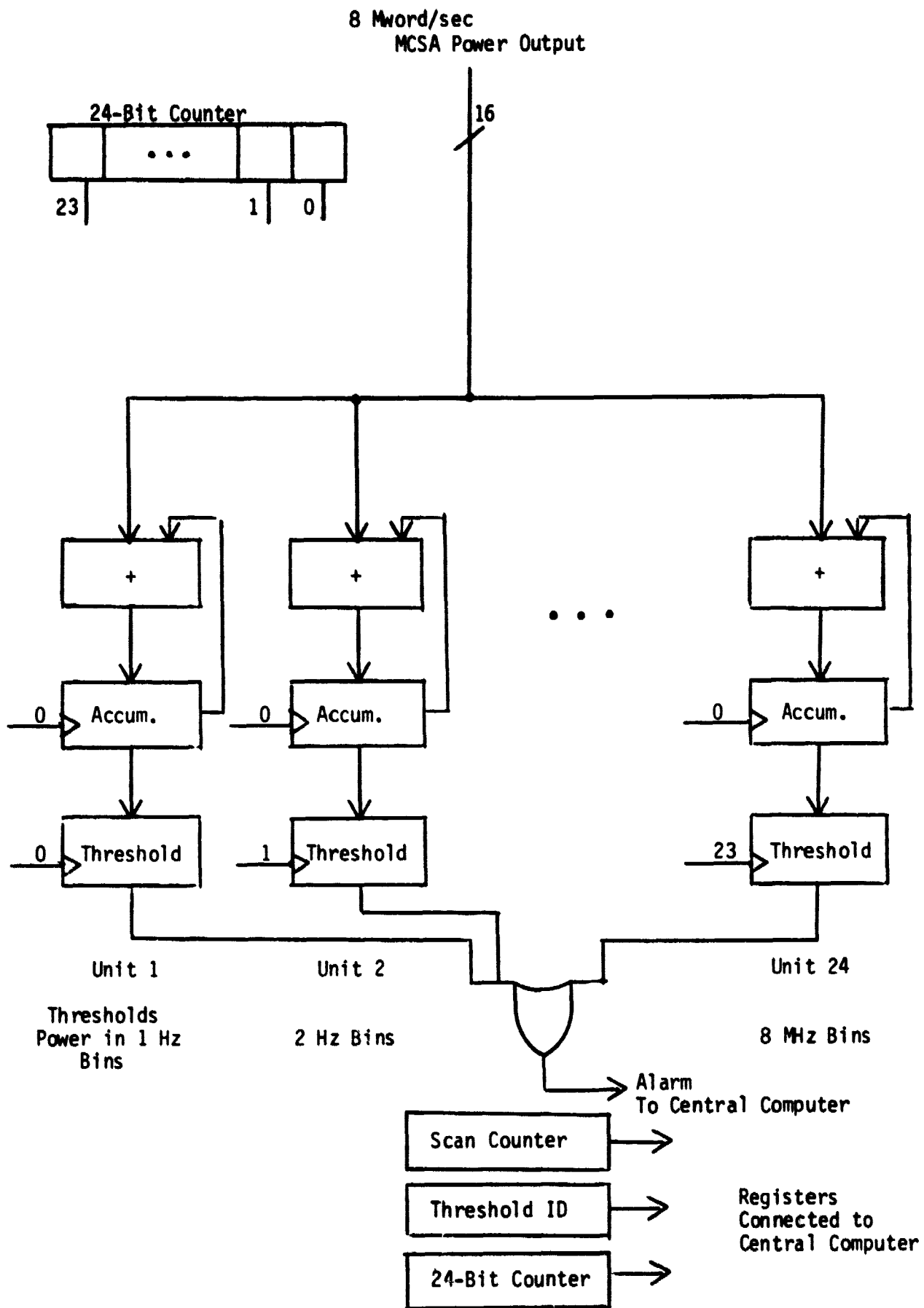


Figure 4.7 - Pulse Detector

integrated scan is read into an identical pulse detection unit (with appropriate thresholds), and a separate alarm, with the associated registers (less the scan ID register), may be communicated to the computer.

#### A Hardware Implementation of the Carrier Wave Detector

As discussed in Chapter 3, this processor is responsible for the detection of negatively and positively drifting carrier waves (drifting in frequency) of drift rate between  $-1$  and  $+1$  Hz/second, and of any initial frequency within the 8 MHz band. The detector integrates for 1000 seconds of data, within which 50 "slabs" of data (measuring 20 seconds by 8 MHz) are individually examined for carriers, as is the grand total of the 1000-second observation. See Figure 4.8.

Through the use of the Generalized Coherence measure, the algorithm provides an enhanced sensitivity to drifting carrier waves. By responding to the phase, polarization, and amplitude or coherence of a signal occupying a moving frequency bin (called a "slice"), the detector improves its signal to noise ratio by a factor of  $\sim 6$  over what is accomplished through simply summing the power within the slice.

The processor also searches for "carriers" of arbitrary bandwidth. One may think of this either as the search for several nearly adjacent drifting carriers, or the search for spectral features of any width (up to 8 MHz) which retain their spectral shape and phase, but with a frequency offset in time. The detection of such drifting signals is accomplished by binning the output of the carrier wave detector across the axis of the starting frequency positions of the slices.

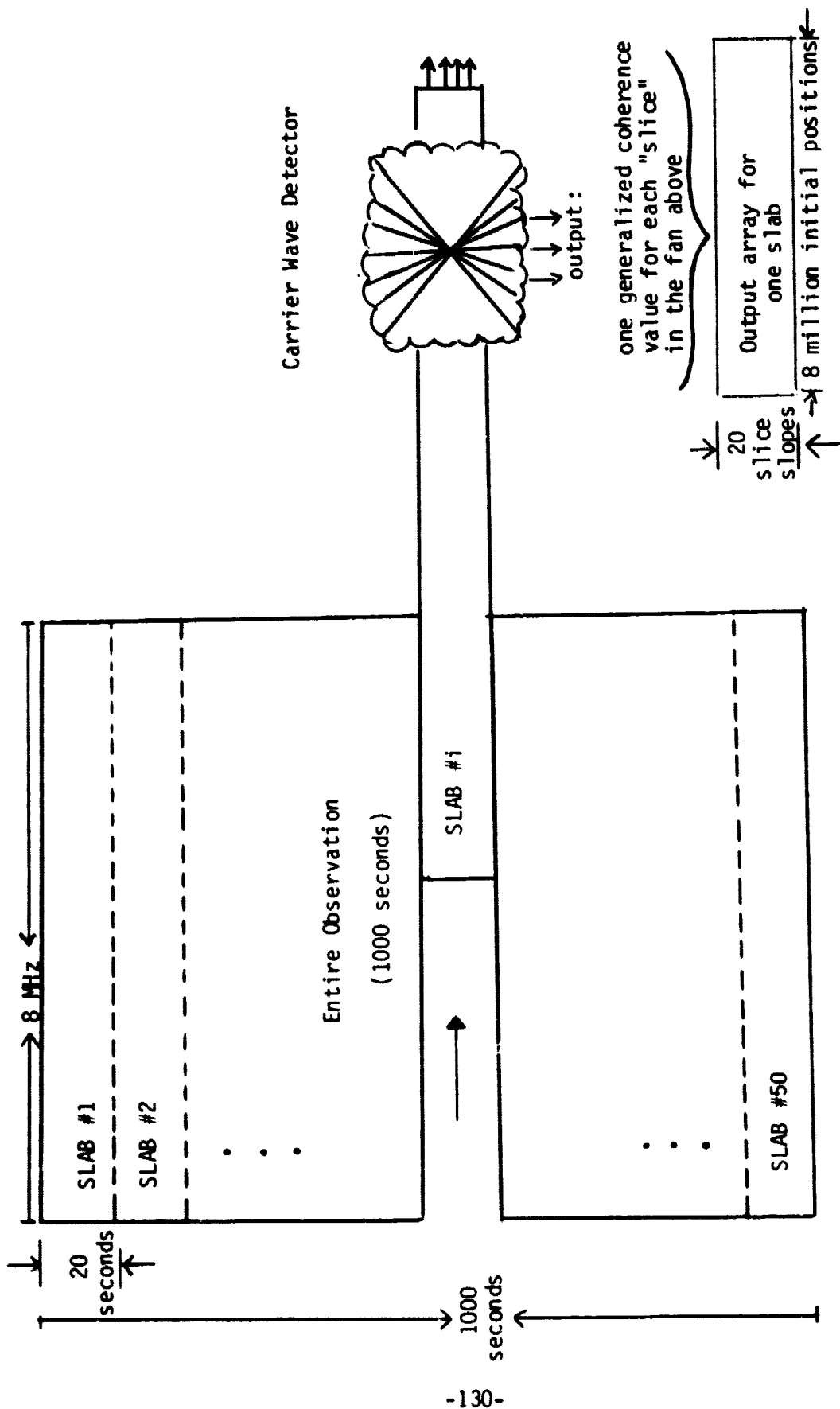


Figure 4. 8 - The parallel approach of the CWD algorithm, where each long horizontal slab is pushed through the "fan" which emanates from a single frequency bin.

## Implementation

The detection of the Generalized Coherence of carriers presents the most difficult processing requirement of the entire OSD, where as many as 200 complex integer multiplies must be done in parallel at an 8 MHz rate, a rate far in excess of what is required of any other processor in the system.

One further ramification of the implementation of this algorithm is the necessity of providing a "horizontal" data flow to the processor as pictured in Figure 4.8. This horizontal flow of data, where data closely associated in time and frequency enter the processor together is accomplished by the array transposer.

Given the somewhat formidable equation for the Generalized Coherence measure of a drifting carrier wave (discussed in detail in Appendix A3.10 and shown on page 136), a complete separation was made of serial operations applied to each slice of data (with a specified initial frequency in the band) and these operations were arranged in a long pipeline with branches. All data necessary to compute one GC value enters the pipeline at an 8 MHz rate in the horizontal manner pictured, and is moved from stage to stage at this same rate.

Since all of the data within a slab (8 MHz x 20 seconds) originates at an 8 MHz rate, all the resulting GC values (8 MHz initial positions x 20 slices) emerge from the pipeline in real time. While the workspace of each stage is relatively small (less than 100 words) RAM buffers are considered the optimal storage medium for their simplicity. The pipeline, with RAM double buffers is shown in Figure 4.9.





The addressing scheme for reading in the appropriate raw data values for a particular slice, as well as the transfer of data from the workspace of one stage to the next is provided by PROM controllers. The first stage controller also maps-in the correct weighting factors for the slice being processed. The pipeline sequentially processes 20 slices for an initial frequency and then begins on the next initial frequency. Therefore the output of the last stage appears as sequential columns of the slice (vertical) by initial frequency (horizontal) array shown in Figure 4.9. As this array appears, it is accumulated onto a 1.3 Gbit disc memory so as to improve the signal to noise ratio of carriers that drift across several, if not all slabs. Some minor address computation is performed for this accumulation process, via the disc (programmable) controller. The controller displaces the data from a particular slice such that, for example, the GCV of a slice drifting from 1 to 20 Hz in the first slab is added to the 21 to 40 Hz slice result of the second slab.

An integration of this type is performed on each incoming slab, so that we have available at all times the most recent slab, and the sum of all slabs to date. This leads to the three tests: 1) the thresholding of each GC value from a slab (160 million values), 2) the coarse binning over 24 different resolutions of a slab, and 3) the coarse binning of the sum of slabs. The coarse binning circuit, as suggested by Figure 4.9, is simply the circuit used in pulse detection (Figure 4.7), only now the input to the circuit is the output array of the CWD algorithm. The result of binning the accumulated output array is shown in Figure 4.10. It should be stressed that the examination of a single slab output over 24 resolutions requires no memory, and, that this test alone constitutes a very sensitive search for drifting signals. Assuming the carrier is

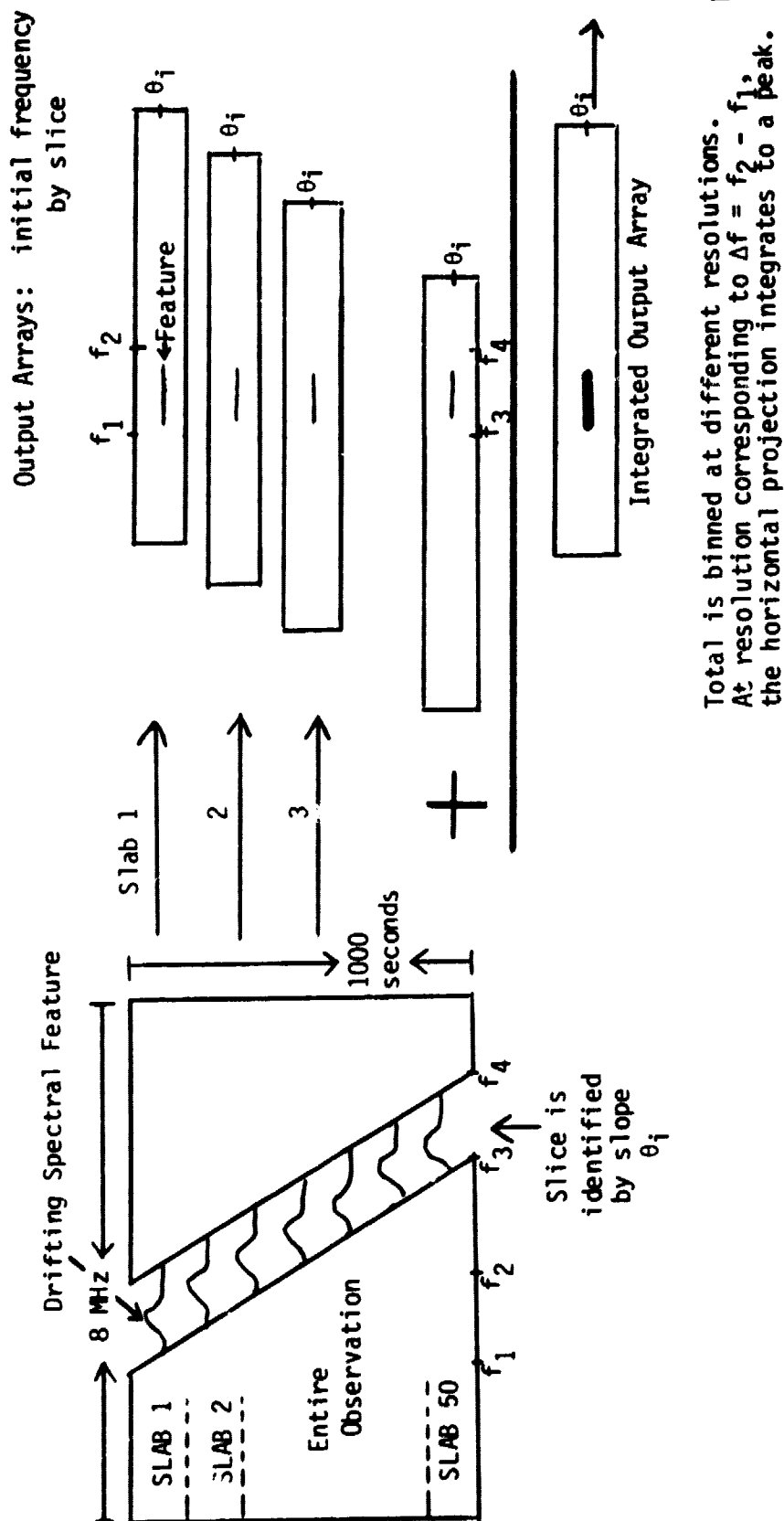


Figure 4.10 - Detection of a broadband, coherent, drifting signal via CMD integration and binning is shown.

observable for 1000 seconds, the 1.3 Gbit disc allows integration which serves to increase the SNR by a factor of  $\sqrt{50} \approx 7$ . For this reason, the optional CWD integration module was placed on the system pipeline (Figure 4.3) below the detector itself.

### The CWD Pipeline

The following schematics (Figures 4.12 - 4.16) show the stages in the pipeline, along with a guide (Figure 4.11) delineating how the branches of the pipelines fork and join. Where there are 20 identical functions occurring in parallel, the diagrams show the outline of 20 boards in a stack. The reason for 20-fold multiplicity is that a single slice traverses 20 scans (rows) of data and the computations performed on the raw data in these rows may be performed in parallel. At some points in the pipeline there are operations which require more than one 8 MHz cycle. These operations then span more than one stage of the pipeline so as to be allotted the required time (e.g. stage A7 - A8 is a 5-stage, 16-bit, cascade full adder). Also, at some points, data is produced by one stage that is not needed until many stages down the line. Such data is passed along the series of RAM workspaces but is shown in the schematic passing through registers marked "H" for hold.

Stages 1 & 2 - The incoming MCSA data are first weighted in accordance with the ray angle, the signal distribution model and position in the array. These factors are predetermined and are reflected in the weighting values stored in the PROM look-up table, designated by  $\alpha$ . Since a ray may cross three frequency bins, three weighting values are multiplied by the three potential bin contributions.

Stages 3 - 14 - These stages compute the various factors which comprise a GCV which is defined as:

$$GCV = \left| -\frac{2m}{(M+1)(N-1)} \left( k_1^2 \sum x_{i+1} x_i^* + k_1 k_2 e^{j\delta} \sum x_{i+1} y_i^* + k_1 k_2 e^{-j\delta} \sum y_{i+1} x_i^* + k_2^2 \sum y_{i+1} y_i^* \right) \right|$$

where  $k_1 = \frac{1}{N} \sum \frac{|x_i|}{\sqrt{|x_i|^2 + |y_i|^2}}$   $k_2 = \frac{1}{N} \sum \frac{|y_i|}{\sqrt{|x_i|^2 + |y_i|^2}}$  as computed in

Stages A3-A9, and

$$|x_i| = \sqrt{R^2 x_i + I^2 x_i} \quad \text{given} \quad x_i = R x_i + j I x_i$$

and  $\delta = \tan^{-1} \frac{\text{Im} \sum_{i=1}^{N-1} x_i y_i^*}{\text{Re} \sum_{i=1}^{N-1} x_i y_i^*}$  (computed at B6-B9)

and  $m = \frac{\sqrt{\sum |x_i|^2 - \sum |y_i|^2 + 4 |\sum x_i y_i^*|^2}}{\sum |x_i|^2 + \sum |y_i|^2}$  (computed in C4-C13).

It is possible, therefore, to compute each of the summations separately, multiply them together and add the results together at the end. The  $k_1$  and  $k_2$  components are derived in Stages A3 - A9,  $m$  is computed in parallel C4 - C11, the factor

$\frac{2m}{(M+1)(N-1)}$  is calculated in C12. Meanwhile, the other five correlation summations are computed in a parallel set of operations in Stages A3 - A9.

The results are combined in Stages 10 - 14.

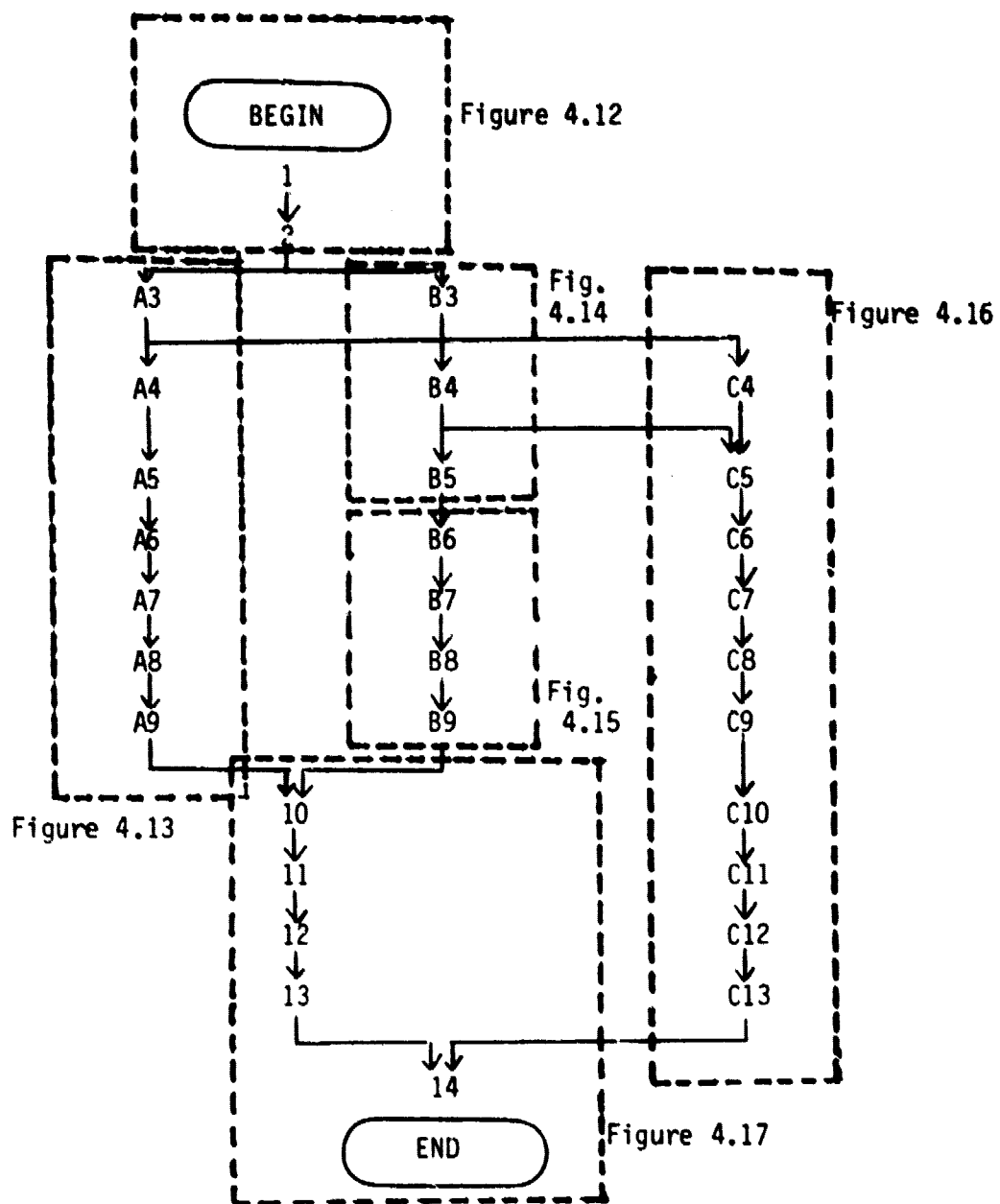
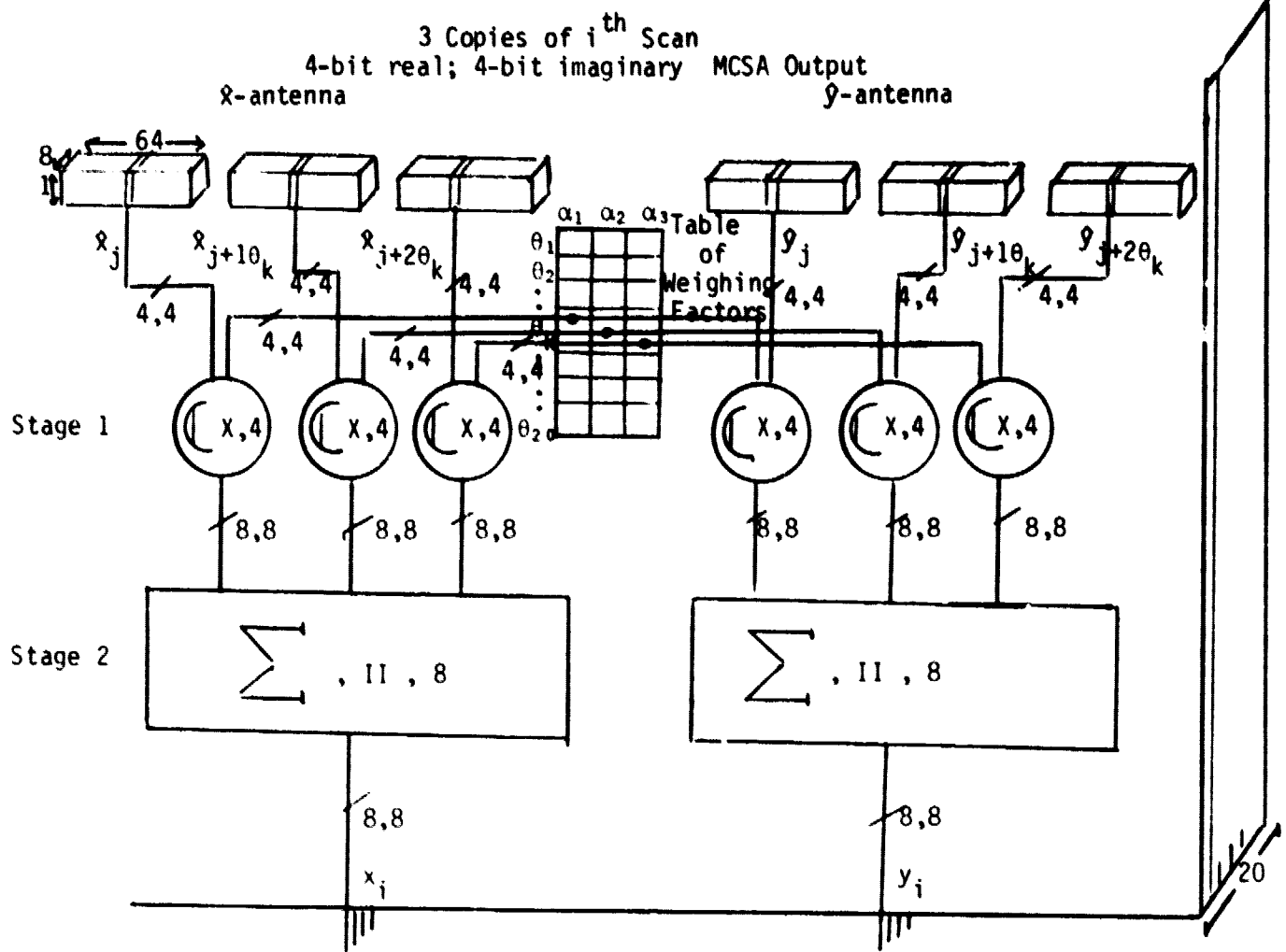


Figure 4.11 - CWD Pipeline



#### Notes on Stages 1 & 2

For each  $\theta_k$  there are three complex (4 bit real, 4 bit imaginary) weighting factors which include an area component and a phase correction component.

The input values are held in the RAM buffers for 2.5  $\mu\text{sec}$ . During that time the ray values for each of the 20 slopes are weighted. One angle is started each 125 nsec.

Not shown is the table which contains the displacements from the  $j^{\text{th}}$  frequency  $10_k, 20_k$  which is used to update the address register for the correct ray factor. The displacement is a function of  $i$  and  $\theta_k$ .

Figure 4.12 - Beginning of CWD Pipeline

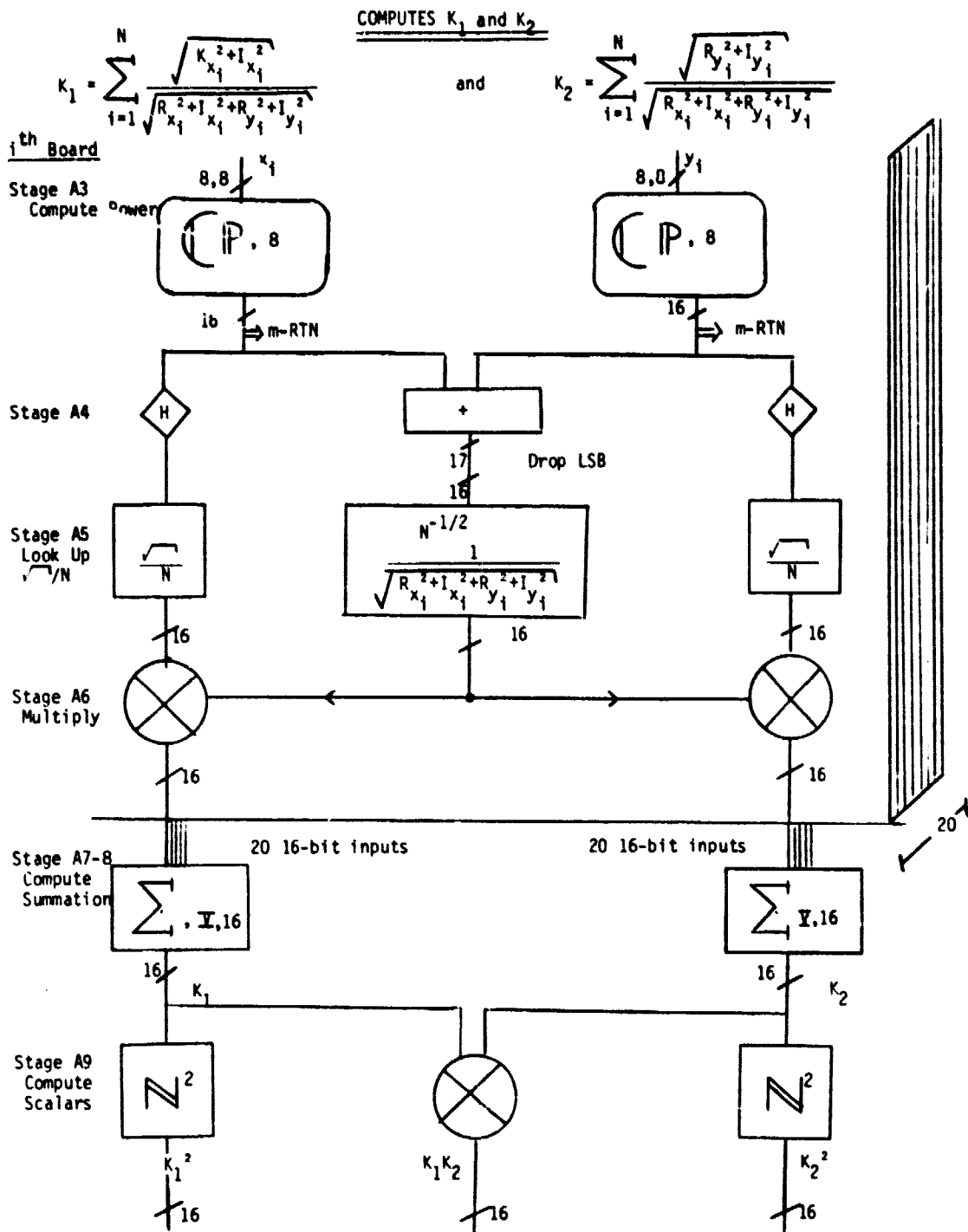


Figure 4.13



COMPUTES AUTO- & CROSS CORRELATIONS

$i^{\text{th}}$  Board  
 $i=1, N-1$

Stage 83  
Compute  
Correlation

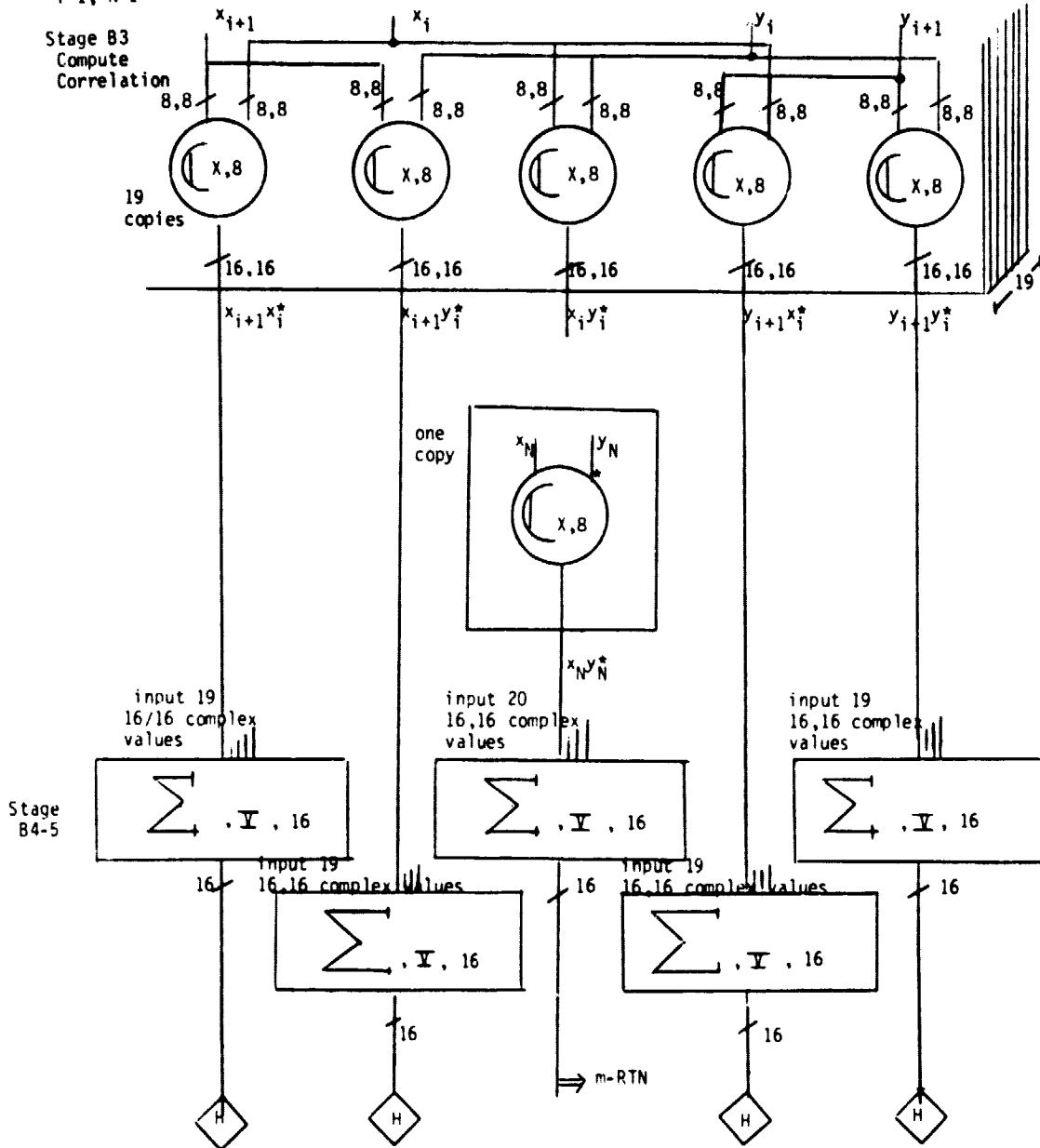


Figure 4.14

51



Figure 4.15

m-ROUTINE: COMPUTES DEGREE OF POLARIZATION

$$m = \frac{\sqrt{(\sum |x_i|^2 - \sum |y_i|^2)^2 + 4 |\sum x_i y_i^*|^2}}{(\sum |x_i|^2 + \sum |y_i|^2)}$$

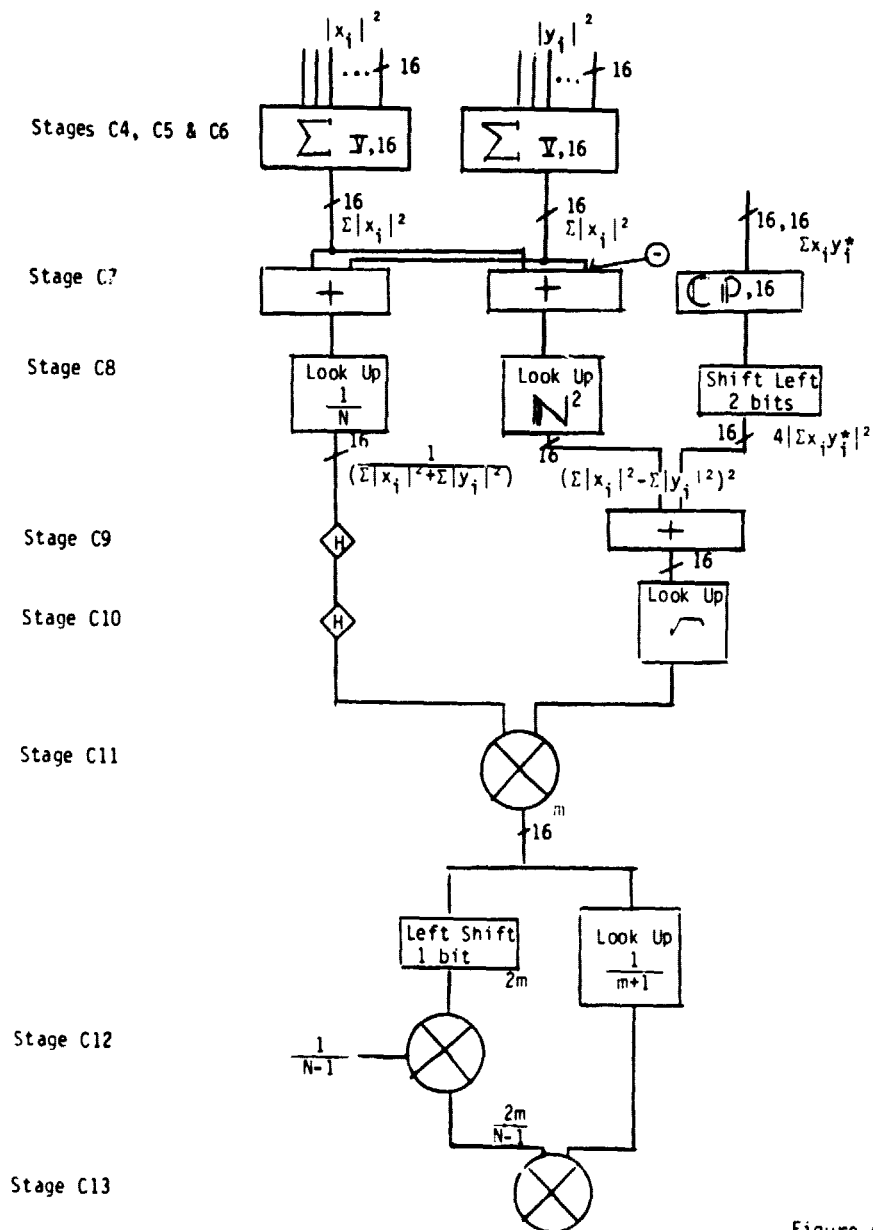
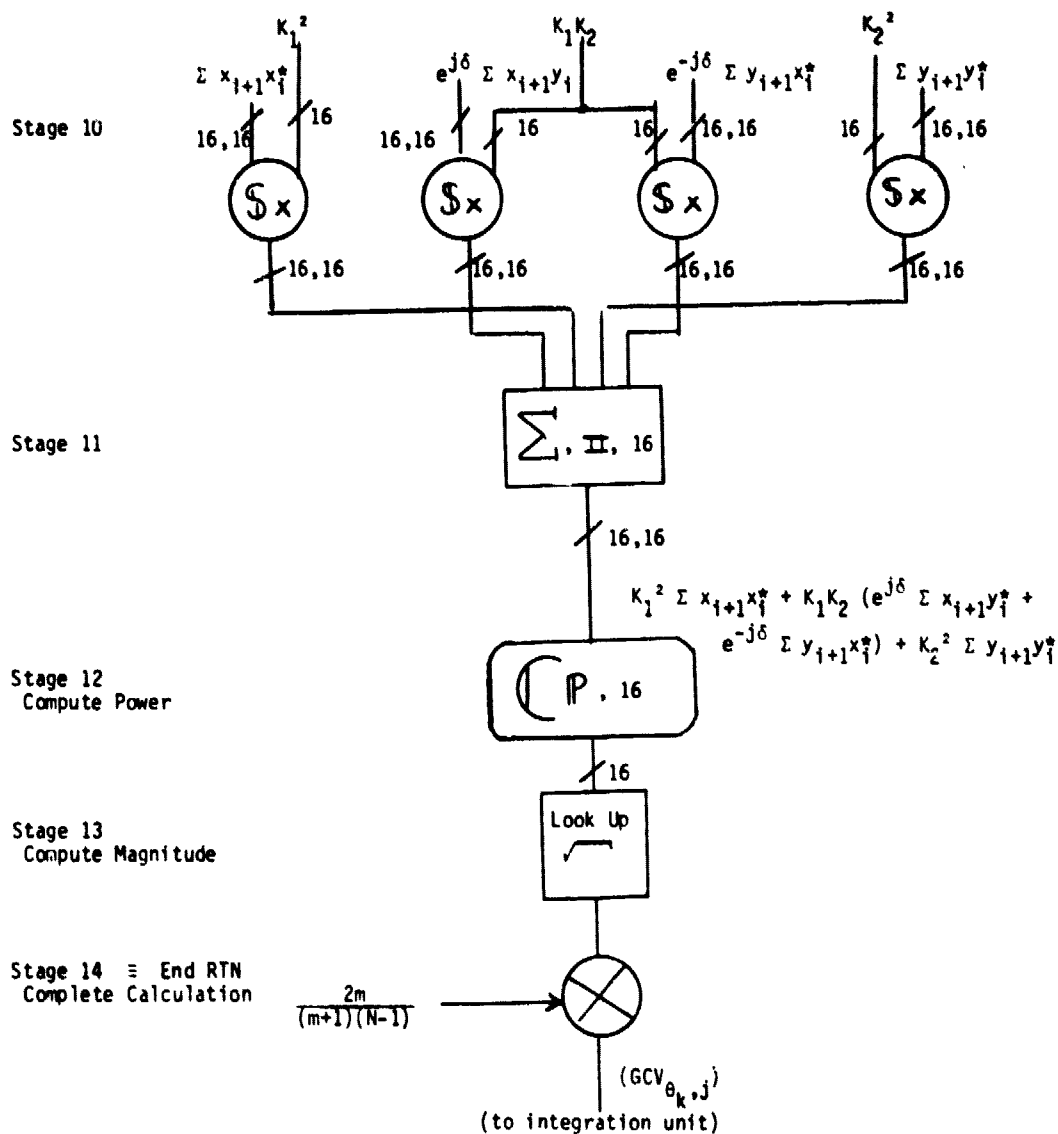


Figure 4.16






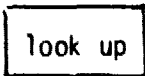
# COMPLETE COMPUTATION FOR $GVC_{\theta_k, j}$



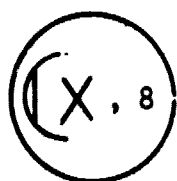
Convert  $\frac{1}{(N-1)}$  to a binary fraction, then use a multiplier and a shift.  
 e.g.  $N = 20$   $\frac{1}{19} = .0526$  and  $\frac{27}{512} = .0527$ .

Figure 4.17

Finally, we examine the actual components of the pipeline. We reduce the complex mathematical symbols shown in the preceding diagrams to real integer adders and multipliers. The key to the symbols is as follows:

	a complex function
	a power computation, e.g. $x^2 + y^2$
X, $\Sigma$	a multiply and addition operation respectively
IV	the circuit involves 4 stages
16	and retains words of 16 bits
	Hold register
	16-bit bus
	change sign bit
	a ROM table lookup of the stored function (e.g. 745189)

The means of performing 8-bit multiplication at 8 MHz rates is easily commercially available (e.g., TRWMPY8HJ circuit, maximum propagation time 65 ns). To do additions at this rate is also no problem (e.g., TTL 745182, maximum propagation time 36 ns). The 16-bit multiplication units commercially available in 1979 are, in the worst case, a little below the required 8 MHz rate (140 ns prog. time = 7.1 MHz), but if this does not improve, there are many alternative means of reaching this rate including multiplexing, or ROM lookups. The approach favored most by the design team is to have custom MSI complex mathematical circuits designed by industry, of the functions shown in Figure 4.18 and 4.19. This should be the most economical solution, and such circuits, being of a fundamental nature, may have commercial appeal in their own right.



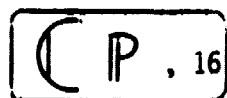
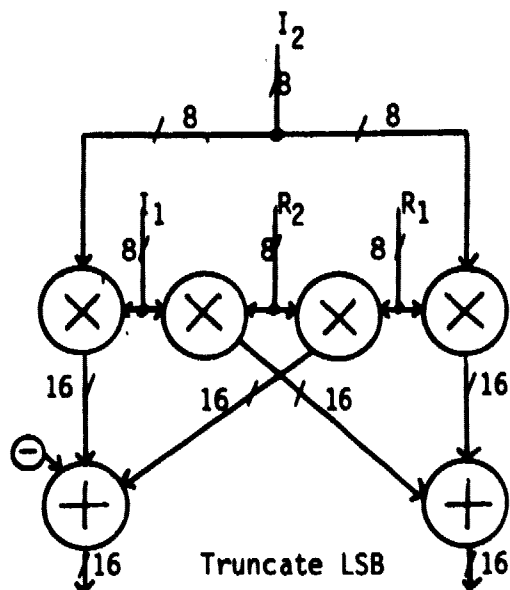
# 8-BIT COMPLEX MULTIPLIER 8+8j

IN:  $A = R_1 + jI_1$  ,  $B = R_2 + jI_2$

OUT:  $\text{Re}(A \cdot B)$  ,  $\text{Im}(A \cdot B)$

MAX RATE: 15 MHz

MAX DELAY: 100 nsec



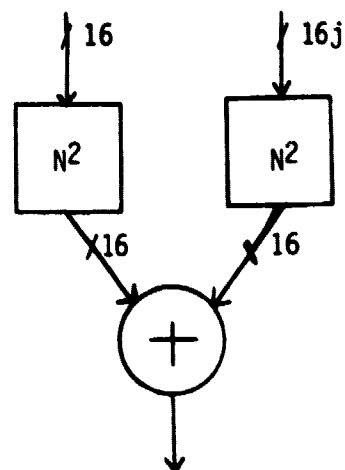
# 16-BIT COMPLEX POWER FUNCTION

IN: one 16+16j number

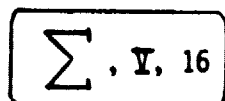
OUT: 16-bit sum of squares

MAX RATE: 20 MHz

MAX DELAY: 55 nsec



# COMPLEX ADDER



IN: 20 (16+16j) numbers

OUT: one 16+16j

MAX RATE: 20 MHz

MAX DELAY: 175 nsec

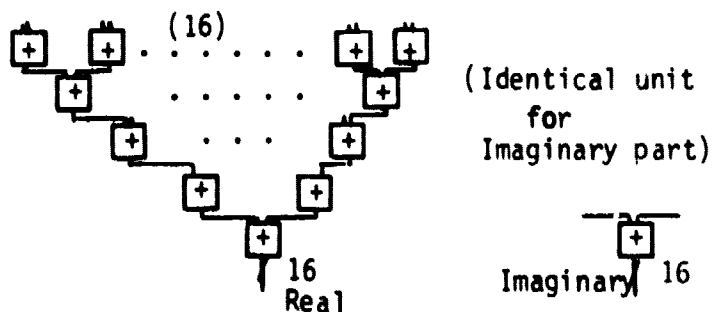
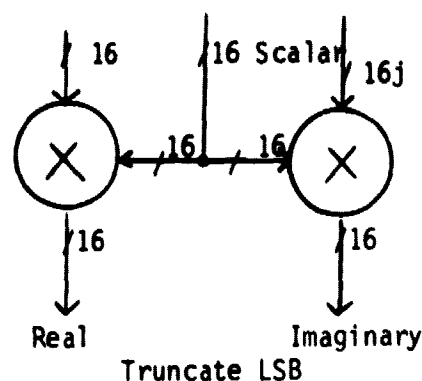


Figure 4.18 - Some of the components used in the proceeding schematics. The values for the rates and delays are derived from current off-the-shelf adders and multipliers.



### SCALAR MULTIPLIER

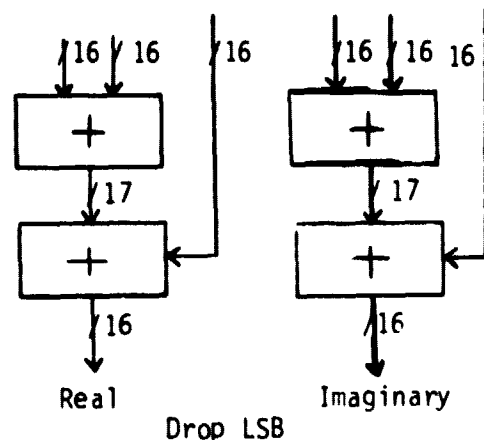
IN: one 16+16j and one 16 bit real  
 OUT: one 16+16j  
 MAX RATE: 15 MHz  
 MAX DELAY: 140 nsec



$\Sigma$  II, 16

### TWO STAGE 16-BIT COMPLEX ADDER

IN: three 16+16j numbers  
 OUT: one 16+16j number  
 MAX RATE: 20 MHz  
 MAX DELAY: 70 nsec



$CP, 8$

### 8-BIT COMPLEX POWER FUNCTION

IN: one 8+8j  
 OUT: 16-bit sum of squares  
 MAX RATE: 15 MHz  
 MAX DELAY: 100 nsec

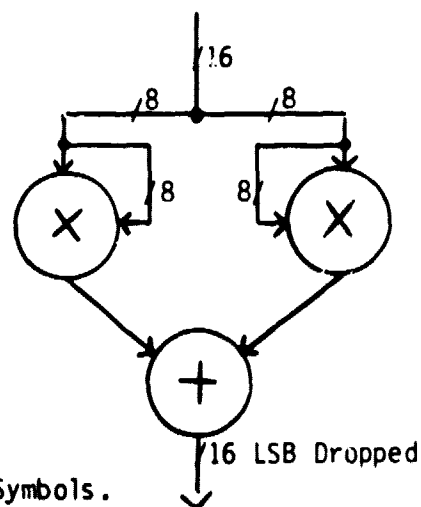


Figure 4.19 - Continuation of Glossary of Circuit Symbols.

## Implementation of the NBIT Algorithms

The implementation of this part of the OSD amounts to the implementation of 9 separate algorithms. It includes, as well, a pre-processor that organizes and prepares the data streams which branch out to the algorithms. Thus, before the data enters the individual processors, it passes through the Overlap System and the Common Processor as shown in Figure 4.20. By the time the MCSA output arrives at the Common Processor, its ordering has changed several times. It emerges from the MCSA scan by scan, from the transpose RAM slab by slab, accessed a column at a time, and finally from the Overlap System block by block, accessed a row at a time. (These terms were defined in section 3.5) The trend here is toward preparing the data to be passed to the algorithms more and more closely associated in the two dimensions of time and frequency, not just in one dimension of frequency.

The blocks contain redundant data. Since blocks overlap 8 Hz on each side with the adjacent blocks, the overall throughput rate is increased to 10 MHz in the system to maintain realtime processing. This is accomplished by the Overlap System. Next, since many of the NBIT algorithms require the same sort input, such as spectral power values, the Common Processor prepares these values at once for many algorithms. Each algorithm takes what is needed from the Common Processor, and then carries through the operation with its own particular hardware configuration.

Before describing the Overlap System, Common Processor, and the algorithm implementations in detail, we will first discuss the role of the NBIT system in the context of the entire system.

It is the NBIT system which compacts the data by a factor of 1000, and reduces the throughput rate from megahertz to kilohertz. It is also the



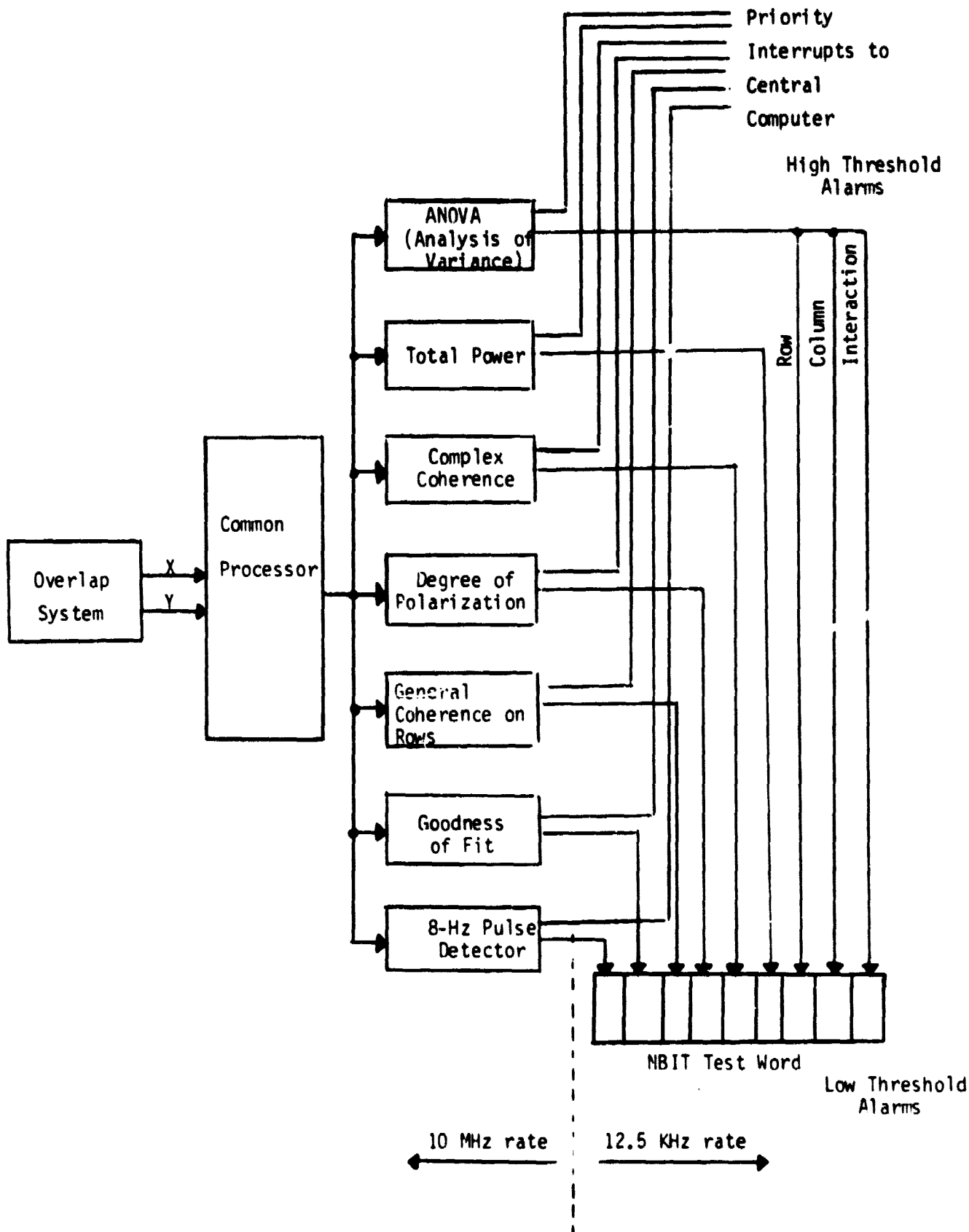


Figure 4.20 - NBIT Detector - Block Diagram

system that surveys the data with high sensitivity, noting millions of candidate low-level signal features, and, rather than bombarding the operator and central computer with these, incorporates them into the compacted array for further scrutiny. Finally, it is a collection of very diverse detection algorithms, a Battery of Independent Tests, which are used together in a "shotgun" approach, with the thought that an arbitrary signal type will be caught by one or more of these tests.

The NBIT detectors are constructed so as to be able to contact the central processor directly (through an interrupt "alarm") if a block of data passes a particular test with a highly significant score. This direct line to the central computer is a common feature in the Oasis System. Each processor has one, and they are designated by triangles in Figure 4.5. This symbol next to the NBIT system indicates there is available to each of the nine processors a high threshold alarm channel to the central computer. (The thresholds for these high level alarms, and the entire strategy regarding false alarm rates and archiving, are discussed in the next chapter).)

The term "low threshold alarm" in Figure 4.20 refers to the fact that each block of 800 numbers is reduced by an NBIT algorithm to a single bit based on a comparison with a threshold. In order to provide the maximum in sensitivity, as explained in Chapter 3, this bit is set whenever the test score for a block is greater than the mean value resulting from a block of noise alone. (Thus, in the absence of signal, the low threshold alarm is triggered by noise half the time.) The collection of low threshold alarms then is not transmitted to the operator, nor to the central computer, but becomes the compacted array, and is stored in Pattern Block Replay section of the Cluster Detector.

Since each of the individual test processors obtains a block (800 numbers) of data at a 10 MHz rate (sped-up to offset the inclusion of redundant data due to overlapping blocks), and each turns a block into a single bit, the bits flow out of the processors at a  $10 \text{ MHz}/800 = 12.5 \text{ kHz}$  rate, into the Cluster Detector.

The Overlap System, the Common Processor, and each of the NBIT processors are relatively simple and memoryless custom-made units, built from the same components as the CWD, though, for the most part, far fewer. In one test (Goodness of Fit) a pair of microprocessors are also used to provide floating point divisions which retain a large dynamic range. Otherwise, the same integer adders, multipliers and small PROM look up tables are employed. The schematics of the Generalized Coherence by Row, and the Degree of Polarization tests are found in Appendix A4.3, and are a much simpler version of that of the CWD processor which has already been presented.

The implementations are memoryless in the following sense. With the array transposer system, we have contrived to arrange the data block by block. so that once a block has been tested, and the low and possibly high alarms are recorded, the working registers of an NBIT test may be completely cleared. Without the array transposer, the data would arrive scan by scan, and partial test values of the 250,000 blocks that comprise 20 scans would have to be stored in accumulators for each test. From an economic viewpoint we note that to provide these (RAM) accumulators would be cheaper then providing the array transpose system. But the transposer is included to make possible the detection of drifting carrier waves, and that it also allows a memoryless NBIT system is simply a side benefit.

To allow the low threshold 9-bit number (representing the results of the nine simultaneous tests on one block) to be synchronously transferred, some pipelining is necessary. Those tests with several stages (e.g., GVC by row) incur a delay in the production of their bit, so that the other quicker tests must provide a pipeline of hold registers to match this delay. Such delay is on the order of milliseconds.

### The Overlap System

As shown in Figure 4.21, data (in slabs, column by column) feed from the 20 Mbit RAM to both the CWD and the overlap RAMs. At all times two of these RAMs are set for reading in at 8 MHz, and the third is writing out a 10 MHz. Each RAM holds 800 16-bit words (40 Hz x 20 seconds). In order to overlap the blocks by 8 Hz on a side, the last 8 columns read into one of the 8 MHz read-in RAMs, are also the first 8 columns of data read into the other. When the first is full it becomes a 10 MHz read-out RAM, and the functions of each of the three RAMs are rotated. The 10 MHz readout is row by row. Not shown, but rather trivial to implement, are the details of the three-phase counter (which counts once each 800 8 MHz cycles, and controls both the Read/Write function of the RAMs, and the clocks selection to the RAMs. Also not shown is the addressing sequence (stored in PROMs) to read out the data in a different order than the sequential readin.

### The Common Processor Implementation

This processor, shown in Figure 4.22, uses previously described components to prepare data streams for the NBIT algorithms. It also aids the "GVC by rows test" by preparing a delayed signal for computing the autocorrelation functions. All preparations are done at a 10 MHz rate in parallel as shown.

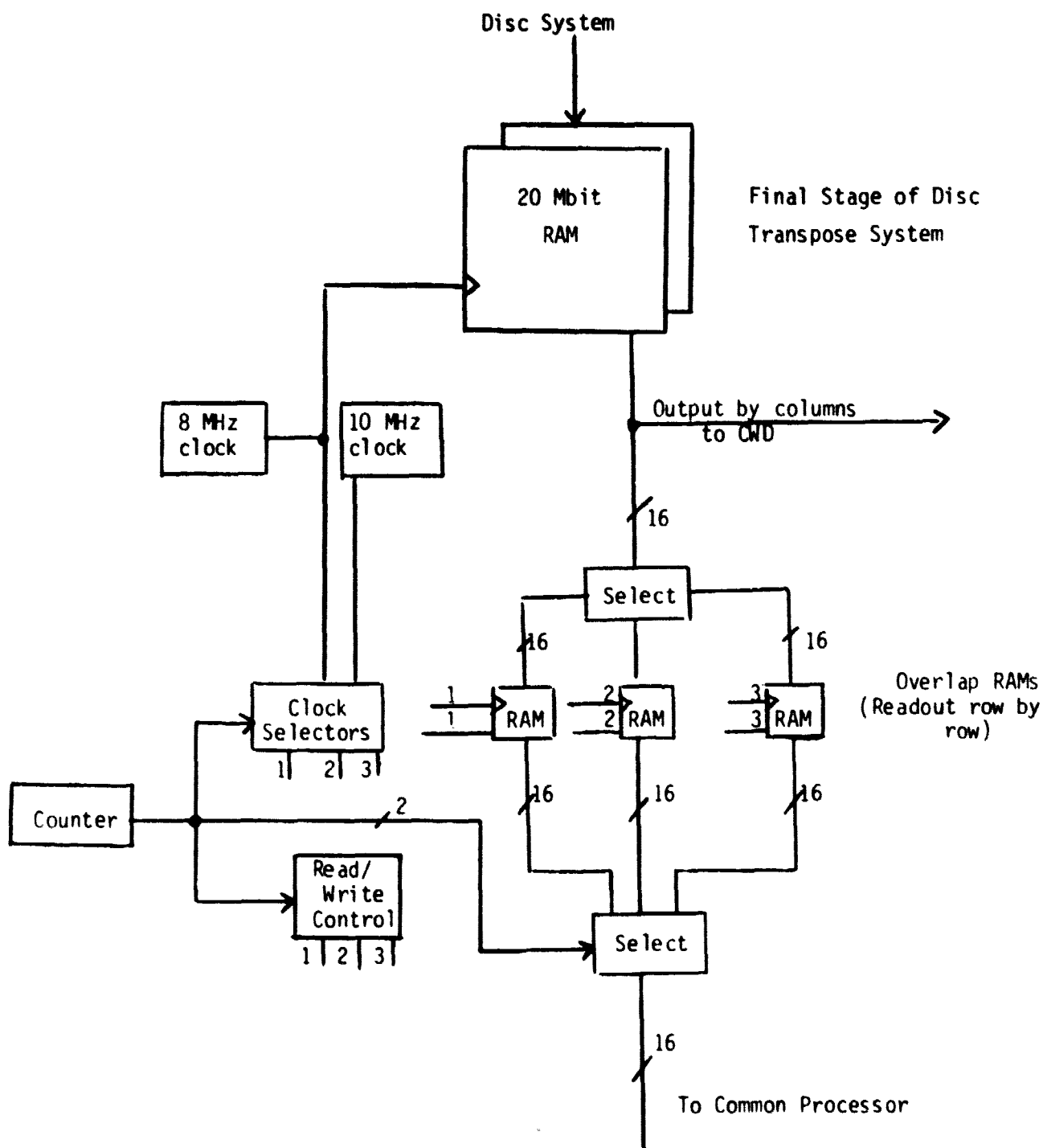


Figure 4.21 - Overlap System. Data is read in to the Overlap RAMs column by column at 8 MHz, and is then read out at 10 MHz, with 25% of the data read out twice.

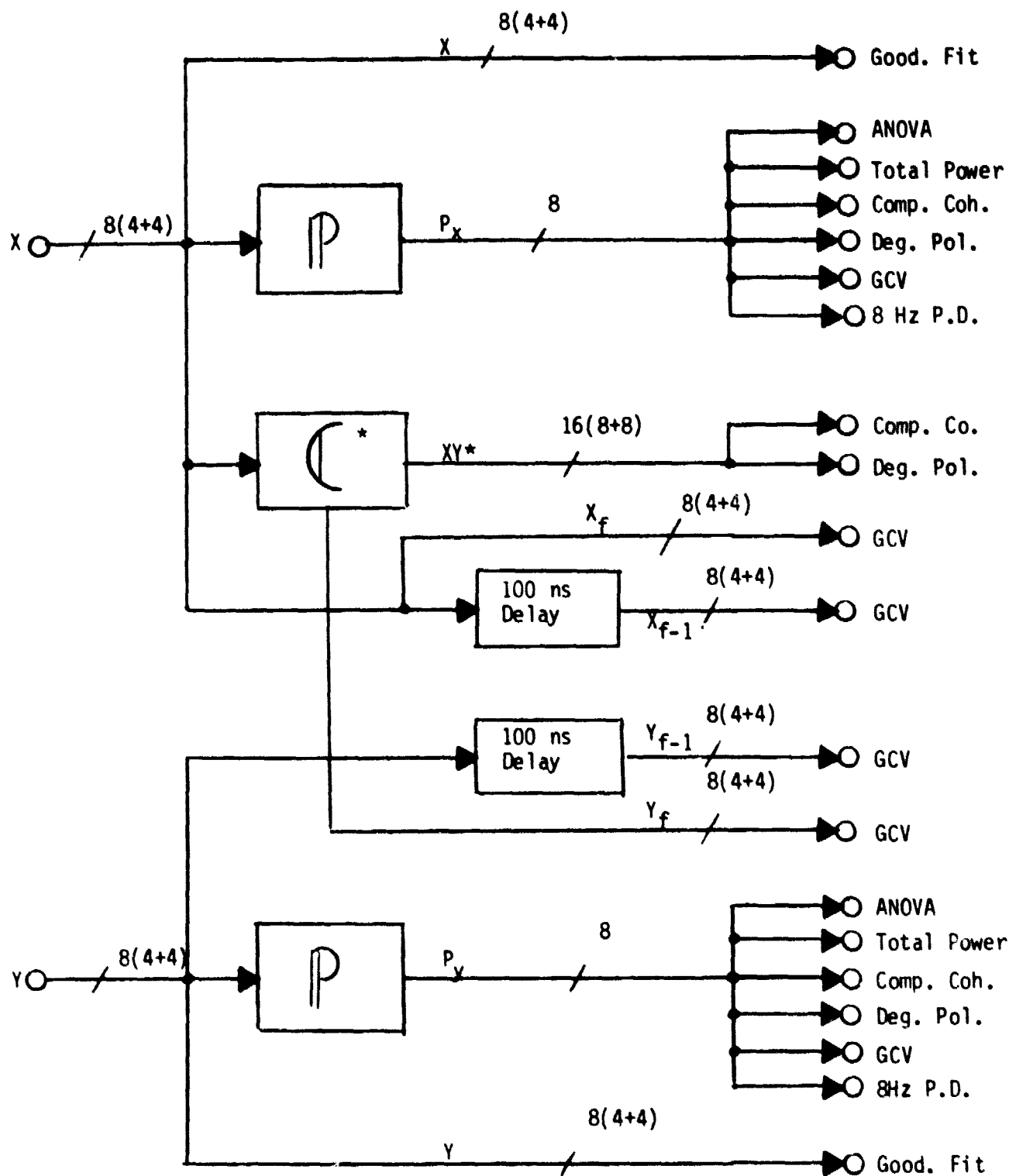


Figure 4.22 - NBIT Detector - First Stage Common Processor

### 8-Hz Pulse Detector Implementation

This detector (Figure 4.23) computes an 8-channel running sum across each row of a block. The total power is compared against a high and low threshold at a 10 MHz rate. Any occurrence of excess power will latch a "1" (low threshold) and will possibly send an interrupt (high threshold). A counter disables the shift register at the start of a new row when less than 8 values are available for summing from the shift register.

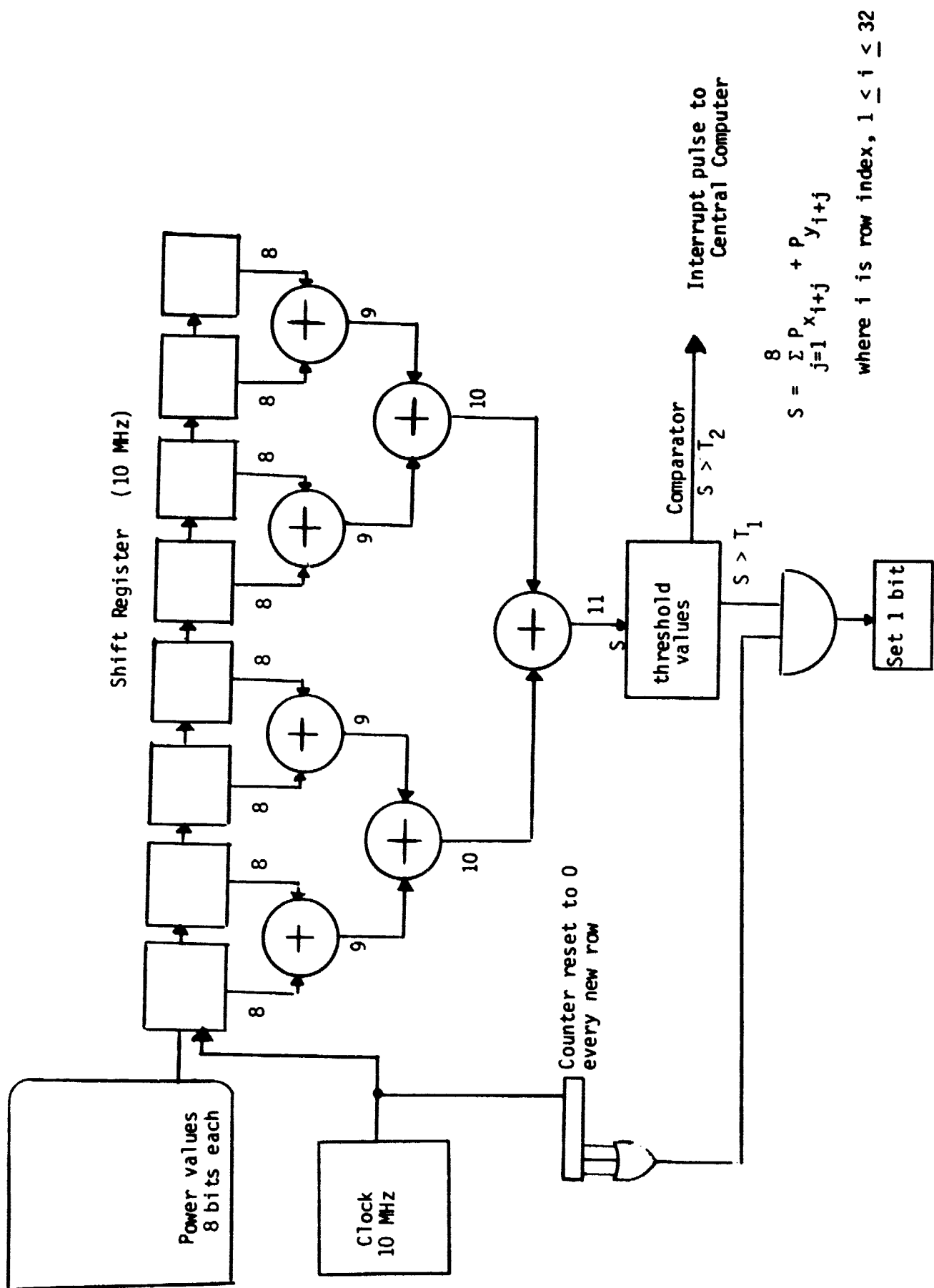
### Total Power Implementation

After 800 accumulations, the total power of the block is compared against the two thresholds (Figure 4.24). Although this is by far the simplest and cheapest detector of the entire Oasis system, it is considered of fundamental importance, as it is sensitive to constant power within a 40 Hz (medium-sized) band.

### Goodness of Fit Implementation

Here the signal's real and imaginary amplitude components are checked for obeying Gaussian statistics. Since, for a 4-bit number, there are 16 quantization levels, these values are decoded (Figure 4.25), and the number of times each quantization level occurs in a block is accumulated in ripple counters. The counting process can occur easily with conventional ripple counters at 20 MHz rate, so that each such unit can handle the real and imaginary part of one polarization.

At the end of the accumulation of the 800 values, a pair of microprocessors perform the Chi-Squared test determining the significance of the deviation of each level. Ripple counter "k" is initialized with  $-T_k$  so that the microprocessor

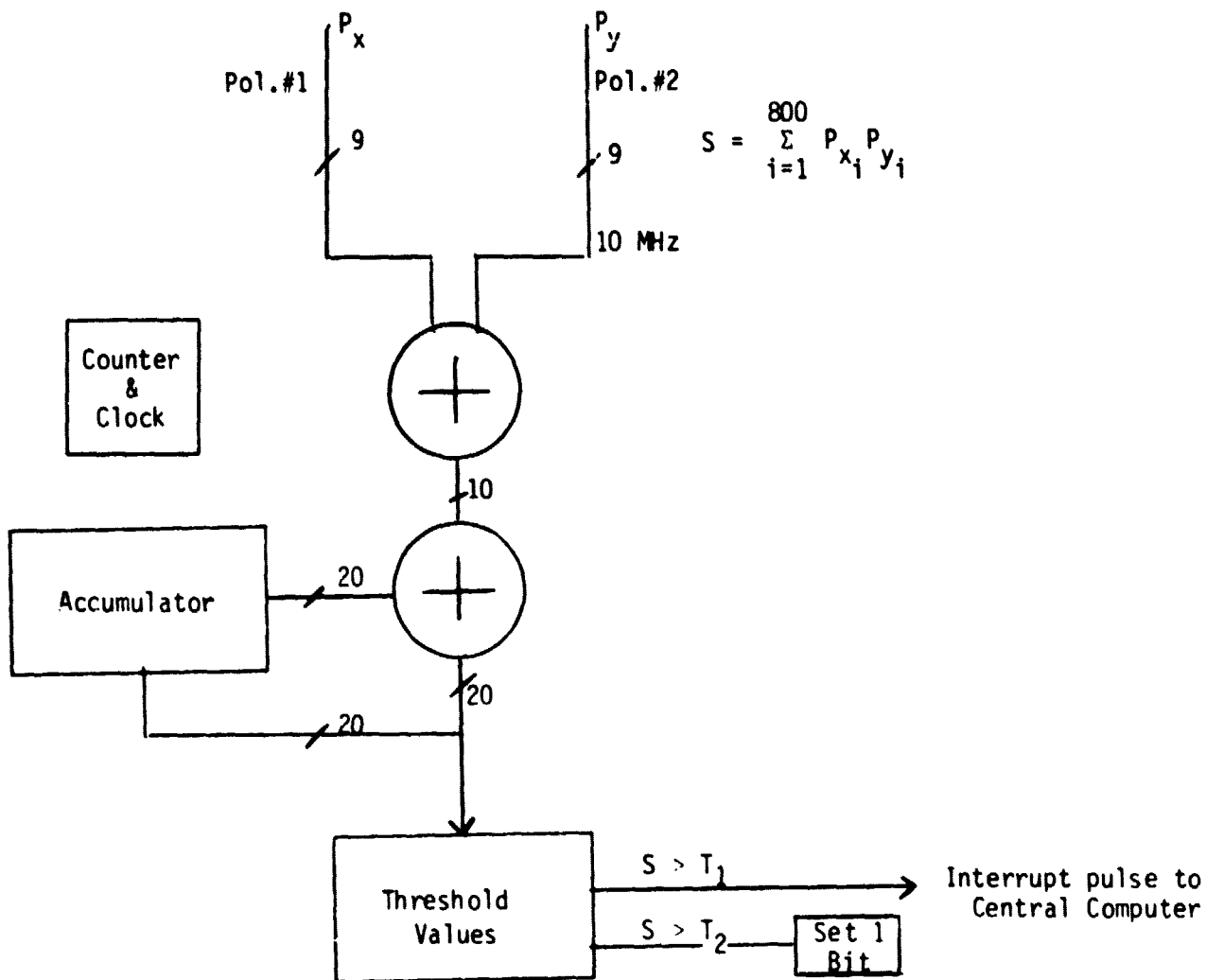


$$S = \sum_{j=1}^8 P_{i+j} + P_{i+j}$$

where  $i$  is row index,  $1 \leq i \leq 32$

Figure 4.23 - 8-Hz (Narrowband) Pulse Detector





Data is accumulated for 800 points (40 Hz x 20 sec block of data) then sent through a thresholding device. The threshold device (magnitude comparator) will decide if either a high-level or low-level alarm has been triggered.

Figure 4.24 - Total Power Detector

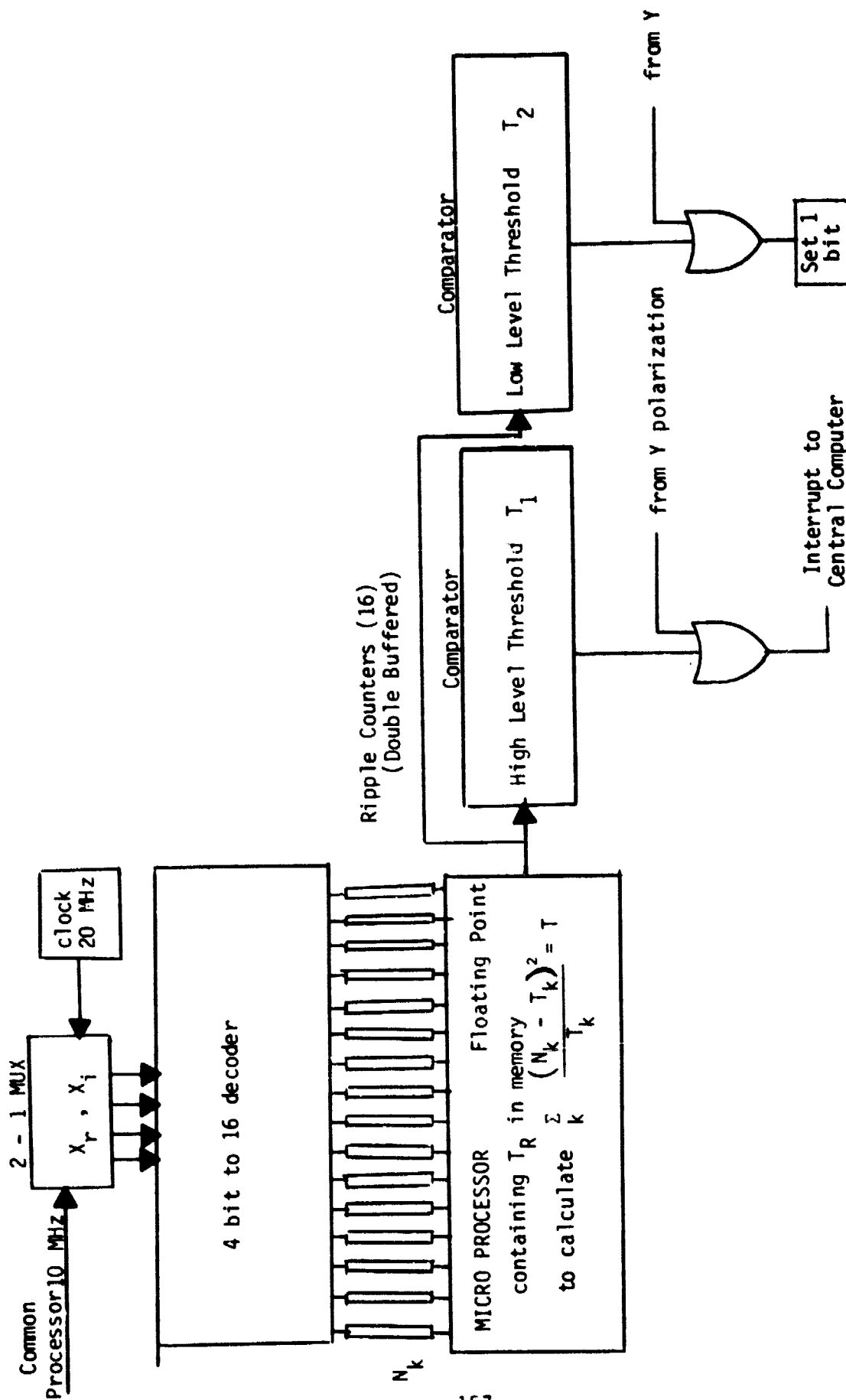


Figure 4.25 - Goodness of Fit for one polarization. The other polarization is handled by an identical unit.

need only perform a square and a divide for each  $k = 1, 16$ . With only  $(10 \text{ MHz})^{-1} \times 800 = 80 \text{ } \mu\text{sec}$  to perform 16 such sequences, a look up table, combining the squaring and division by a constant operation into one read might be required to provide the sufficient speed.

### Complex Coherence Implementation

The Common Processor provides this detector (Figure 4.26) with  $\text{Re}(XY^*)$ ,  $\text{Im}(XY^*)$ ,  $P_X$ ,  $P_Y$ , so that only accumulators and look up tables are required to easily implement this algorithm.

### ANOVA Implementation

The implementation of the ANOVA algorithm computes the equations listed in Table A3.8.1. Ideally, the analysis of variance used in this study (a two-way ANOVA cross classification) requires a matrix of  $R$  rows and  $C$  columns in which two statistically independent sets of measurements of the same target are contained in each cell of the matrix. Because time is a factor in the analysis, to simultaneously obtain a second set of independent data would require a second antenna of the same polarization (not simply a second polarization).

A suitable substitute for a set of independent replicates has been shown to be a staggering of each second's output of the MCSA. The first second's scan fills the first half of the first row of power values, while the second second's provides the replicates. Similarly, the third and fourth seconds' scans comprise the second row of the matrix. Despite the displacement in time of these signals, simulations have shown this approach to be acceptably effective (see A3.9, Simulation Studies).

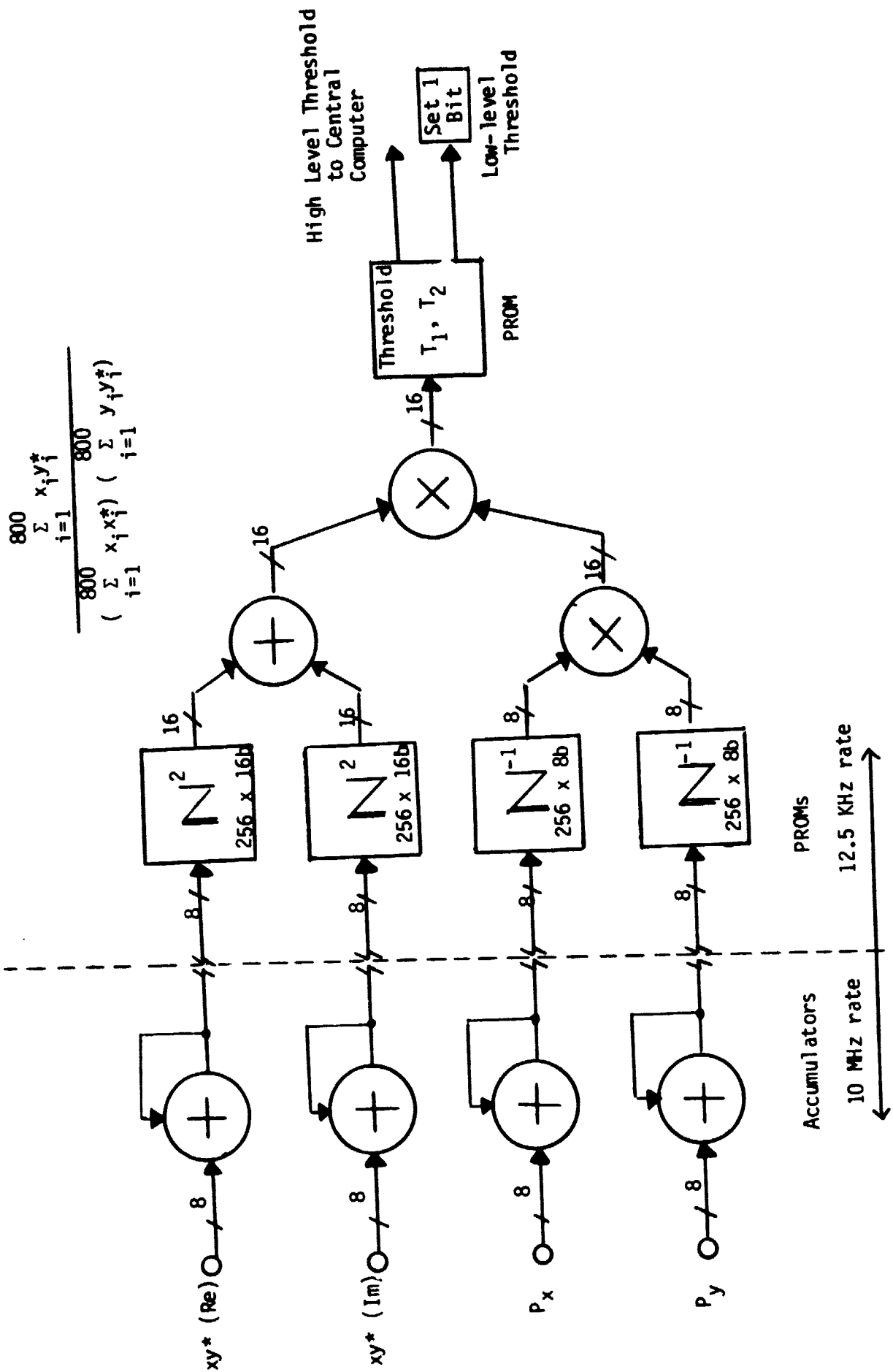


Figure 4.26 - Complex Coherence Detector. Most of the computing has already been done by the Common Processor.

As shown in the table, the quantities MSR, MSC, MSI and MSE require computation. It is not difficult to expand these four summations so as to obtain:

$$MSE = \frac{1}{rc} (A - \frac{1}{n} E)$$

$$MSR = \frac{1}{nc(r-1)} (D - \frac{1}{r} B^2)$$

$$MSC = \frac{1}{nr(c-1)} (D - \frac{1}{c} B^2)$$

$$MSI = \frac{1}{n(r-1)(c-1)} (E - \frac{1}{c} C - \frac{1}{r} D + \frac{1}{rc} B^2)$$

where A, B, C, D, and E are defined in Figure 4.27. By expressing the MS values in this way, the entire computation may be performed in the parallel pipelined manner shown. Here, we have also taken advantage of the dual availability of row and column outputs of each block provided by the NBIT common processor. The data flow rate in Figure 4.27 is divided down at each accumulation point indicated, while the bit width of the busses and accumulators is increased.

At the conclusion of each block input, the values A, B, C, D, and E are synchronously buffered from their respective accumulators into the micro-processor which computes:

$$F_R = \frac{MSR}{MSE} = \frac{5 C - 1/10 B^2}{9 A - 1/2 E}$$

$$F_C = \frac{MSC}{MSE} = \frac{20 D - 1/40 B^2}{39 A - 1/2 E}$$

$$F_I = \frac{MSI}{MSE} = \frac{200 E - 1/20 C - 1/10 D + 1/400 B^2}{351 A - 1/2 E}$$

where we have substituted from above, setting  $r = 10$ ,  $c = 40$ , and  $n = 2$ . A simple comparison circuit then compares the results against high and low thresholds, and, after oring with the results of an identical circuit for the orthogonal polarization, deposits the three bits in the NBIT word.

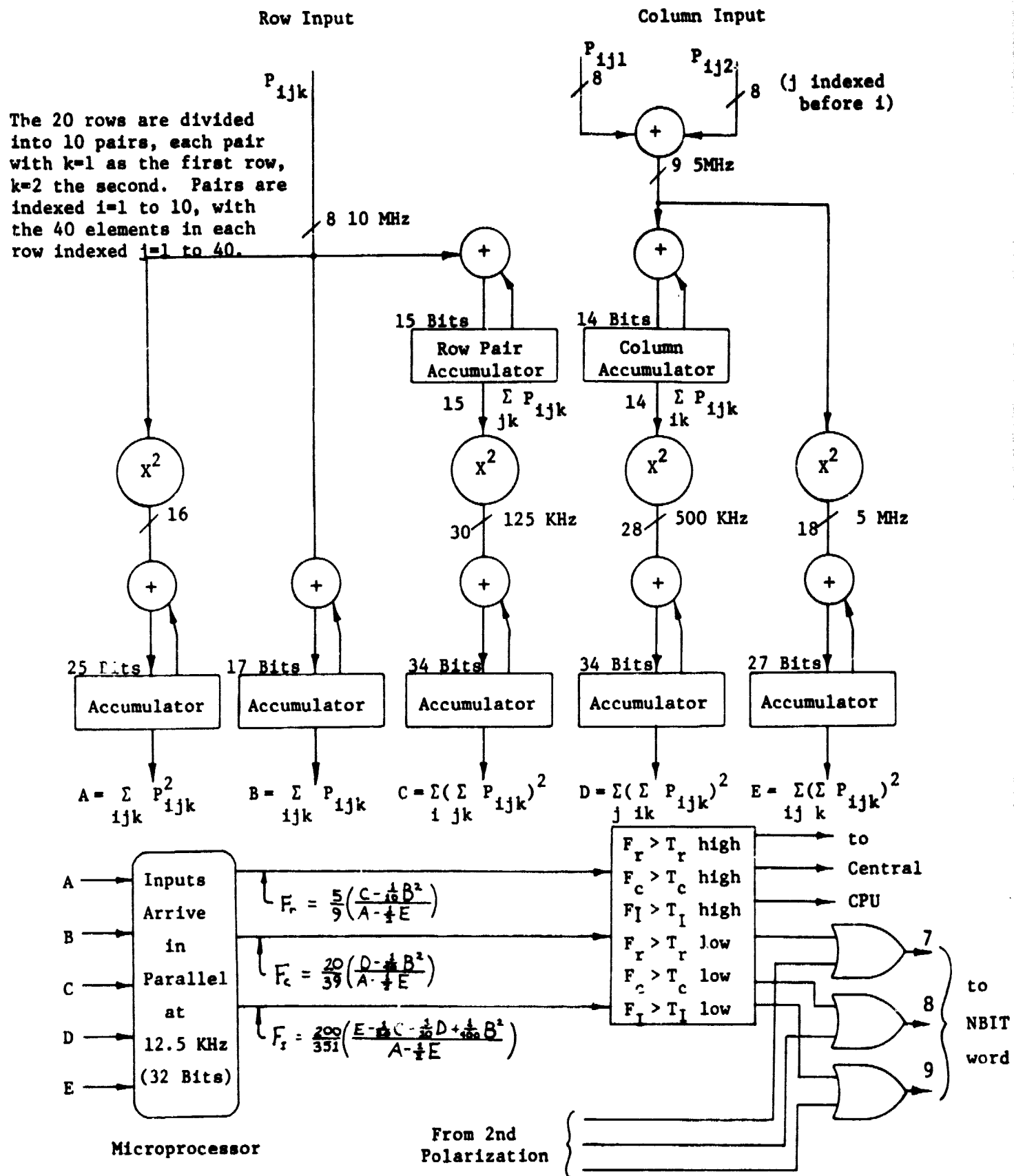


Figure 4.27 ANOVA Implementation showing hardware for one polarization. (Second polarization implementation is identical.)

## Cluster Detector Implementation

The reduction in data rates (8 Mword/sec + 12.5 Kword/sec) allows for sophisticated, non-determinist processing at low cost. We can now use the speed and versatility of a computer to comb through a condensed version of the entire observation. The histogramming (checking for excesses of low-threshold bits) takes place as the observation proceeds, and as each 9-bit value is output for NBIT. At the end of the observation the histograms are completed and the cluster detection (via microprocessor) proceeds. The significant results of this algorithm are transmitted to the Central Computer during the observation of the new target.

The steps of the process are described in detail in Appendix A3.11, and are summarized below.

### The Histogram Circuit

This circuit prepares a "histogram" for each of the 1250 pattern blocks comprising the entire array. The histogram is a number count of the 9-bit outputs of NBIT, according to occurrence of each output type ("pattern class"). There are 512 such types, and these types are not mutually exclusive. Each output of NBIT is thus subjected to 512 tests, and the implementation for these tests is shown in Figure 4.28. A 9-bit number appears for  $(12.5 \text{ kHz})^{-1} = 80 \text{ usec}$  and 512 "mask tests" are conducted at  $512 \times 12.5 \text{ kHz} = 6.4 \text{ MHz}$  rate. The test is for the presence of 1's, with the other bits being "don't care" bits. The "don't care" bits are represented by 0's in each mask, so that the Boolean operation is  $(\overline{M}_1 \vee N_1) \wedge (\overline{M}_2 \vee N_2) \wedge \dots \wedge (\overline{M}_9 \vee N_9)$  resulting in True (1) whenever all 1's in the mask M appear as 1's in the NBIT number N. The clock runs at  $2 \times 6.4 \text{ MHz} = 12.8 \text{ MHz}$  so as to give the histogram RAM both a read and write cycle for each accumulation. The histogram circuit, due to the nature of the input to the NBIT processor, (a row of blocks

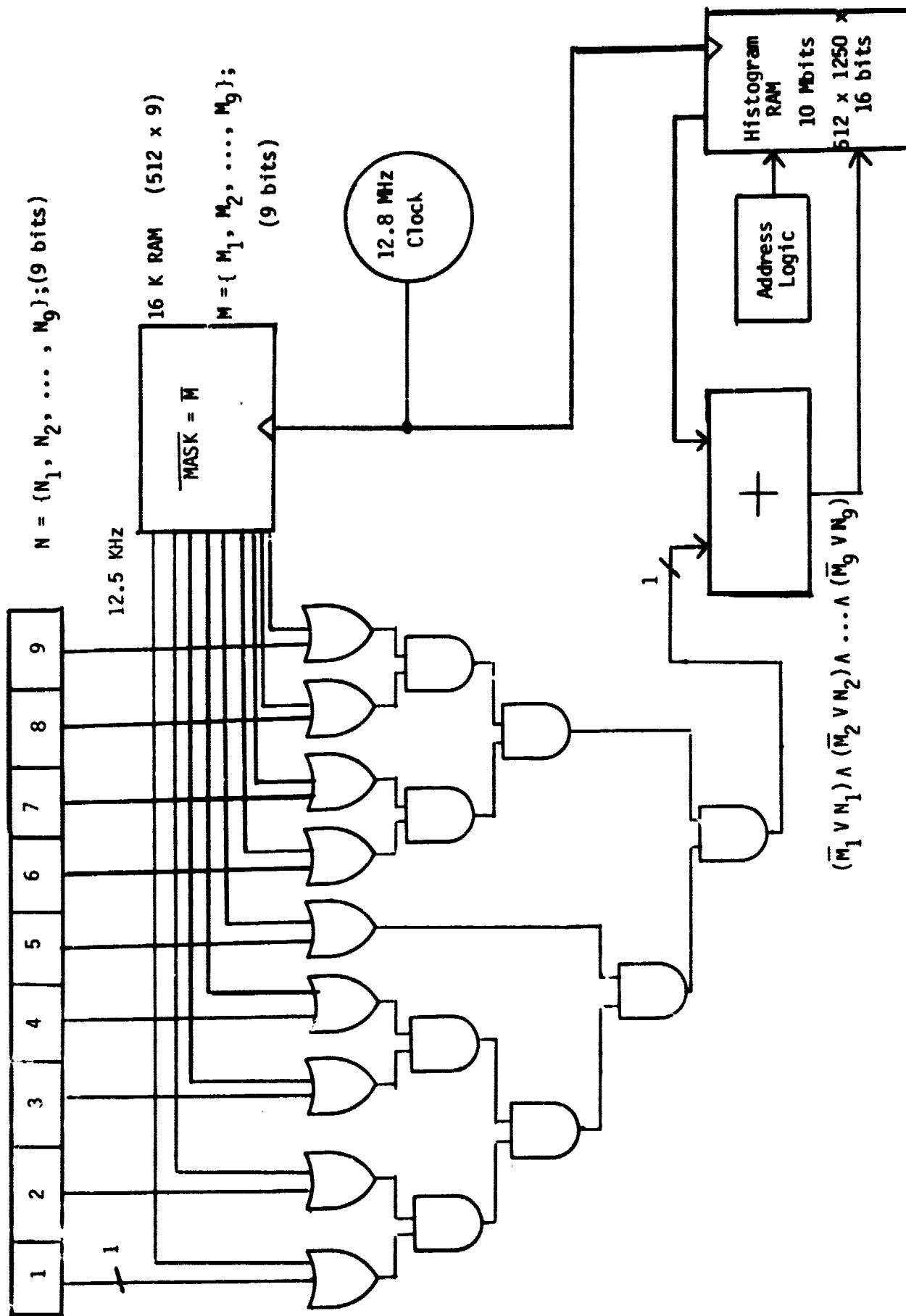


Figure 4.28 - Histogram Generator.  $N$ , the NBIT output, appears each 12.5 kHz, and in that interval is tested against 512 masks from PROM, each test capable of incrementing that bin of the histogram for that pattern block.



at a time) processes a row of all pattern block histograms at a time. Thus the address logic in Figure 4.28 performs a read and write by rows of consecutive pattern blocks.

As the pattern points (NBIT numbers) arrive at the CD system, they are simultaneously stored in an additional memory shown in Figure 4.5 as the Pattern Block Replay. These 4 RAMs constitute an identical pair of double buffers. One pair is for the storage constantly taking place during an observation, and access by the CD processor. The second pair is for storage and access, but this time, by the operator, who looks at interesting pattern blocks found by the CD processor and Histogram Generator and stores these blocks on tape.

#### Cluster Detector Microprocessor

In this section of the CD system, control is transferred to a microprocessor. For simplicity, the system includes only one processor, but it will be shown here how, at this stage in the data flow, large gains can be bought at a very small price by including additional processors conducting nearly identical operations to the one detailed.

The tasks of the microprocessor are threefold:

- 1) Accessing histogram results; sending high-level alarms, and using results to establish a potential pattern.
- 2) "Cleaning" the pattern blocks to be searched.
- 3) Subjecting the pattern block to the adjacency algorithm.

The first task above occurs once at the start of the detection process. The microprocessor has the list of expected histogram number counts that would be found from noise alone (shown in A3.11). It notes pattern blocks possessing significant excesses. If the count is highly significant, a high-level alarm

is sent to the Central Computer. Otherwise, the most significant 1000 excesses are noted in a list in the microprocessor's memory. (It retains 1000 "slots" in memory, and as the histogram contents are read out, the least significant values from either the slots, or from the current read-out is discarded in a "quick sort" fashion). At the end of the histogram read-out, the microprocess has a list of 1000 pattern types, each related to a particular pattern block, that it will perform tasks #2 and #3 on. The sorting of the top 1000 such pairs should take:

1250	histograms
x512	bins per histogram
x~10	slots to search through till minimum is found
<u>x 5 <math>\mu</math></u>	seconds per compare and $\Delta$ card
32	seconds

On the average, only ~10 slots need be checked before a histogram bin value is discarded, because we are selecting only the most interesting 1000 pattern classes out of 1250 x 512 candidates, so that most histogram bins will quickly be determined to be uninteresting (non-significant).

Since, as shown above, we have used 32 seconds out of 1000 seconds available, 968 seconds remain to "clean" the blocks and find any patterns.

The cleaning operation amounts to arranging a bit (cell) pattern in the microprocessor's memory that contains 1's wherever the interesting (excessive) pattern class occurs within the pattern block, and a 0 where it does not. Thus, in the microprocessor's memory we retain the two dimensional position of the "on" cells.

Such a program was written in a high-level language (BASIC) simulation to produce Figures 3.13-16 and its low-level language equivalent might be:

			<u>μsec</u>
loop 10,000 times	→ BACK:	LDM, PNTR, 1	get pattern point 2
		ORI, MASK, 1	is it in class? 2
		ISZ, PNTR, 1	yes, leave cell on 1
		STI, PNTR, #0	no, off, go to next 2
		BRZ, PNTR, OUT	done?, leave 2
		HOP, BACK	1
			<u>10 μsec total</u>
OUT:			

Using a mock-assembly code we have conservatively estimated the run-time of a microprocessor to clean an entire pattern block is  $10,000 \times 10 \text{ } \mu\text{sec} = .01 \text{ sec}$ . Thus, after cleaning each of 1000 blocks we still are left with most of 1 second to search each for patterns.

As a good example of the firmware programming needed, we choose the most time-consuming (with half the cells on by noise) adjacency algorithm; one that checks and counts nearest cell neighbors for complete occupancy.

A cell's survival to the next generation depends on it being totally surrounded. The surviving cells are counted, with a large count indicating the presence of a cluster. This would correspond, in the notation of A3.11, to requiring 8 cells be on within a distance of 1 cell.

Again, such a computer program was used to produce Figure 3.16. The mock-assembly language equivalent would be:

						<u>μsec</u>	
→ START:	LDM,	PNTR,1	get cell			2	
10,000 times	IBZ,	PNTR,START	cell empty? get next			1	
	CLR,	I	set up index			1	
NEIGHBORS:	TRR,	PNTR,PNTR2	get centered on cell			.5	
	ADD,	PNTR2,UP(I)+	move up,down, or none			1	
	ADD,	PNTR2,RIGHT(I)+	move right,left,or none			1	
	ISN,	@PNTR2	neighbor on? skip			1	
8 times ↑	BR,	START	off, get next cell			1	
	TBRB,	2,I,DONE	8th neighbor? done			.5	
	BR,	NEIGHBORS	no, get next			1	
DONE:	INC	COUNTER	record survivor			1	
	ISZ	PNTR,START	next cell?			2	
ALL DONE:						Total	62 μsec
	.						
	.						
RIGHT:	-1,	-1,	-1,	0,	0,	1,	1
UP:	-200,	0,	200,	-200,	200,	-200,	0 -200

We can conservatively traverse the neighbors of each cell in a cleaned pattern block in 62  $\mu$ sec, so that the number of such surviving cells in a pattern block could be counted (COUNTER) in  $62\mu \times 10,000 \text{ cells} = .62 \text{ seconds}$ , and thus in 620 seconds for all blocks. So, with a 1000-second budget, we can find 1000 interesting pattern-point excesses, clean those blocks, and count the cells surviving an adjacency search in 662 seconds. Those excesses are then relayed in a high-level alarm to the Central Computer. The operator, if he wishes, can view such pattern blocks on his image processor, or can automatically let the most significant ones--i.e., with the most survivors--be automatically stored on the Special Interest Archive.

The algorithm just presented is good at finding blobs, or dense amorphous clusters. It can be seen from careful study of Figure 3.16 that extended sparse contiguous structures may be revealed through repeated, more relaxed survivor criteria. If we allow the case of many survivors through relaxed adjacency requirements in one pass, a second pass requiring survivors to have surviving neighbors would then select such patterns. This would require the inclusion of more microprocessors, and would gain the ability to detect filamentary structures within the pattern blocks.

In conclusion, to quote one member of the group, "signal detection and identification hinges on a good pattern recognition scheme". While the implementation here represents a somewhat basic technique, the foundation for easy computer processing has been provided, and, at this stage of the pipeline, any number of sophisticated software schemes could be brought to bear at little additional expense.

## Conclusions Concerning Realtime Processing

This completes discussion of the realtime processing system. All units requiring special hardware or software have been fully laid out so as to demonstrate their feasibility. Options have been given here and in A4.2 regarding implementations of all or part of the proposed system. Problematical aspects of the design have, in all cases (we believe), been explored and solved rather than omitted or ignored. Those details that were omitted are of the banal sort, well within the domain of conventional digital design.

### 4.3 - Archive System Implementation

An essential part of the observing system consists of the ability to play back an entire 1000-scan observation of a star<sup>5</sup> during the observation of the next star or possibly at some later time as an archive. This requires two high-speed recording systems, double-buffered for simultaneous recording and play back, as well as archiving.

The most inexpensive systems capable of handling the data rate (128 Mbits/sec) and the storage requirements (.1 terabits/star) were found to be high-speed helical tape drives. The optimal 1982 version of one of these drives has a storage capacity of 1 terabit. This is rather useful in that it eliminates the need for tape changes for every new star being observed. As many as 20 stars may be recorded before one observation must be rewritten over, or else the tape be changed. The decision to rewrite new data over old data is actually tantamount

---

5 The processing of 1000 scans takes place during the taking of the same scans, and continues on during the taking of the next 1000 scans. It is assumed for simplicity in this study that each 1000-scans corresponds to the total observation of one star. Both shorter observations, and longer observations (1000 scans followed by a second observation) could be easily accommodated into the present scenario, in 20 second increments.

to the decision to erase data or to archive it. Since one tape drive stands available for operator perusal during each new observation, information from the alarms that were triggered during the realtime processing may be used to re-examine the raw data during the 16 minutes = 1000 seconds the tape drive is available. At the end of this period, a new source is observed and the tape drives switch function. If the decision to save the data rather than write over it is made, then the new data is recorded further along the tape, leaving the data from the previous star intact. After 10 stars are archived on one tape, the tape must be removed and a blank one spooled up. Commercial specs for these drives that will suit the need are mentioned in Appendix A4.1.

The total observation archive stores its data block by block (again, as a benefit of the array transposer). This enables the operator, when replaying blocks on this sequential access device, to rapidly view and store adjacent such related data, rather than waiting for the drive to seek and read out each row of a block separately. This is of great advantage, for this sort of high bit read-out drive has long start and stop times (Appendix A4.1) and the operator has only 16 minutes to decide whether an observation should be permanently archived or not.

#### 4.4 - Operator Interface

##### Central Computer

The overall operation of the Oasis signal detector is controlled by the software operating system of the central computer. In this section we discuss its major functions and specify the minicomputer configuration that carries them out.

### Threshold Setting

Each test in the detector yields a value which is then compared against a threshold. The threshold is optimized for the most feasible balance between false alarms and missed signals. Threshold setting can be carried out in either of two modes. In the automatic mode, the CPU polls the response of the various tests while the receiver is observing a quiet patch of sky, or a white noise source. Under these conditions, the sensitivity of a particular test can be adjusted so that noise does not cause an undesirable number of false alarms. These levels can be adjusted regularly during the search while the antenna is "off target". Alternatively, the operator may desire to manually instruct the CPU to increment or decrement a particular threshold as desired. Each threshold is stored at a particular address and communicated to tests via an I/O port.

The threshold setting function of the CPU provides a means of system self-checking and calibration. In the event a part of the system malfunctions, the CPU can be of utility. For example, if a particular test consistently triggers when only noise is present, the CPU will disable that test and notify the operator.

### Detection Handling

After thresholds are exceeded and alarms (or "detection flags") are set, the signal is sent in the form of a priority "interrupt" and the CPU is charged with responding to those. The detector setting the interrupt sets an identifying bit in a CPU register and transfers other relevant data via an I/O interface.

At this point the CPU consults its RFI catalog for possible information bearing on the detection. The catalog consists of two classifications of

signals. The first is "natural" signals which have already been cataloged in various radio astronomy observations, and previously-known RFI; frequency, amplitude, time of day and/or position bounds for these phenomena are recorded. Also, the local RFI receiver dynamically updates the catalog. Local RFI signals are saved for 25 seconds, long enough for the rest of the system to utilize them. The maintenance of and referencing to this catalog represent a considerable portion of the CPU's real time responsibilities.

After a detection is compared against the RFI catalog and is found to not match any previously-known signals, nor have been concurrently received by the RFI antenna, the CPU performs one or more of the following:

1. Alerts the operator (only if desired).
2. Informs the display CPU to make the block available for audio/visual scrutiny. This can also be done manually, by the operator.
3. Prepares for the automatic storage of the block onto the special interest archive, when the total observation archive is free.
4. Records information on the detection in the system log.
5. Flags total observation tape record for saving.

### System Log

The CPU maintains a system log, on standard magnetic tape, which includes information defining parameters of the search in progress, system status, and all significant events and detections.

### Self-Diagnostics

Each subsystem of the Oasis detector possesses a diagnostic processor capable of verifying correct operation by supplying special test data and examining the computed results. Diagnostics are performed during intervals



between star searches on command by the central CPU. Each diagnostic processor notifies the central CPU of the result of its test; the aggregate of the tests is reported to the operator via a display of the system block diagram.

The diagnostic processors are microprocessors with programs in ROM. In some cases, test data are also stored in ROM. In others, test data are generated by the diagnostic programs or by special hardware devices as, for example, with the array transposer diagnostic systems.

The Oasis system operation will in general be degraded but not precluded by the failure of one or more subsystems. The usual result will be to render one or more detection flags meaningless. The central CPU or operator can inhibit the setting of such flags simply by setting thresholds high.

Verification of the control CPU itself must of course be performed at a different time. Diagnostic routines for the CPU, its I/O devices, and the graphics system, are supplied by their respective manufacturers.

#### Other Functions

Other system service functions the CPU performs include: initializing, resetting, on/off control of the system. As part of the initialization the CPU can inform interested algorithms of the length of the ensuing observation. These functions also have manual overrides.

#### RFI Receiver and Catalog

This subsystem consists of log spiral and vertical antenna, a preamplifier, and a remotely-programmable spectrum analyzer interfaced with the Oasis control CPU. The spectrum analyzer is capable of locating spectral peaks or sweeping

the spectrum with a variety of resolutions and rates. It is used by the control CPU to maintain a dynamic catalog of probable RFI signals. A commercially-available analyzer with 1024 channels is used for this purpose. Other RFI data can be entered into the catalog from tape (via the system log unit) or manually by the operator.

The catalog in its complete form is stored on the catalog disc. In addition to frequency, any available information on amplitudes, times of day, antenna coordinates, etc., are included. An abbreviated catalog of frequencies only (for quick checking) is in the control CPU memory.

Entries are static or dynamic. The former are entered via magnetic tape. Dynamic entries are subject to review and removal by the RFI receiver.

The central CPU consults the catalog in the course of processing detections, to determine validity. It has the capability of compensating for any Doppler correction applied to the observatory local oscillator. It can also cause the spectrum analyzer to display for the operator portions of the spectrum surrounding any specific detection.

#### 4.5 - Radio Astronomy System

The Radio Astronomer is given his own station. The output of the coarse or fine resolution stages of the MCSA are made available to this system so as to provide bandwidth for varying applications. Note that Stokes parameters are computed as shown in Figures 4.30 and 4.31. The use of this information is discussed in Chapter 5. The observer also has a 256-Mbit RAM to enable quick subtraction and comparison of observations. Standard observatories (depending on location) may make up the post-processing backend shown in Figure 4.29.

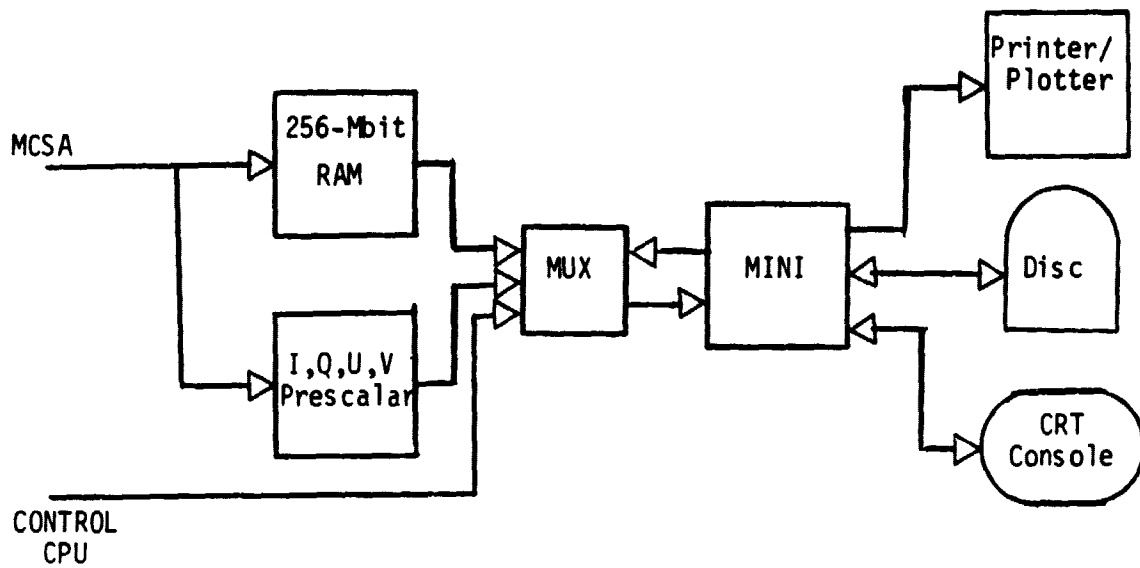


Figure 4.29 - Radio Astronomy Subsystem

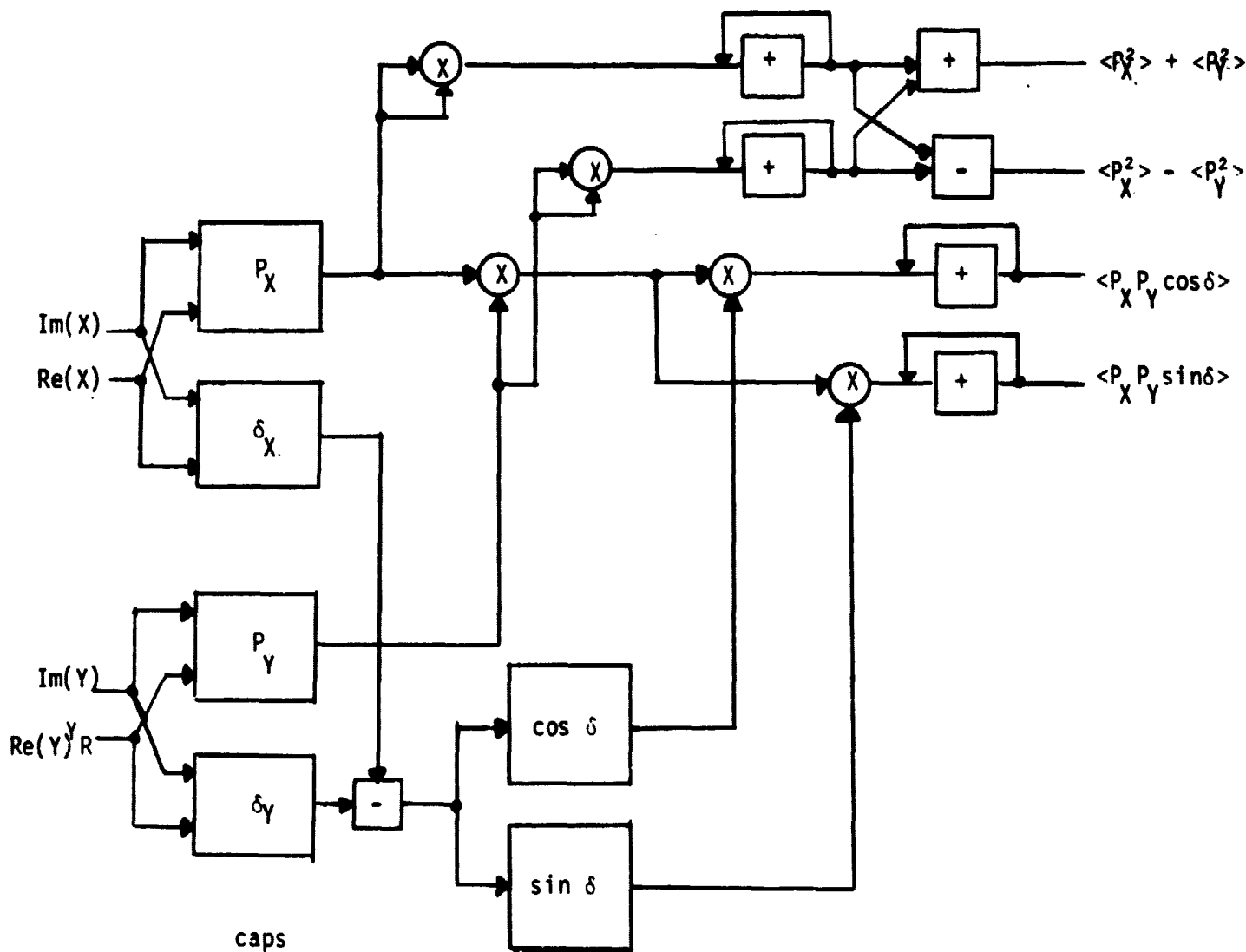


Figure 4.30 - Prescalar ( is computed in the manner shown in figure 4.15)

## 4.6 - Gross Attributes

### Storage Capacity: 2 Terabits Online

At all times the Oasis System has a 2-terabit online data store. This database contains the latest 20 seconds of observations (stored on disc), as well as any previous 4-1/2 hours of observations on two high speed high density helical tape transports. Such transports are currently commercially available today (Ampex "Super HBR"), with each tape holding one terabit. Included with this tape drive, the Oasis system will support a tape library where 20 tapes will store the observations of hundreds of stars, at .1 terabit per star (1000 seconds of observations).

Since the Oasis system is primarily a real time processor with the vast majority (95-99 %) of all input data discarded for lack of interesting features, it is expected that an entire stellar observation would be archived in the tape library only for exceptional reasons. Therefore the system supports a collection of smaller random access memories for intermediate computation and processing. There are approximately ten independent "working" memories in the system, each a solid state dynamic RAM, with total storage capability of 1.1 gigabits.

In summary the system supports:

- 2 terabits, Tape, Online
- 40 terabits, Tape, Offline
- 7.8 gigabits, Disc, Online
- ~1 gigabit, RAM, Dedicated

The role of these memories was delineated in the description of memory stream preparation as well as in the discussion of the individual algorithms and archives.

## Brief Inventory of Major Components

The following list is intended to give the reader a feel for the variety and scale of the hardware required to implement the system.

- 2 Super High Bit Rate (HBR) Tape Transports
- 4 1.3 Gbit Parallel Transfer Disc Drives (removable heads)
- 1 Minicomputer/Display Processor Unit
- 1 Minicomputer, referred to as "Central Computer" in this report
- 7 Special Purpose Hardware Processors - TTL at 10 MHz clock rates, varying in complexity from 1 to 10's of integrer multiplies and/or adds per cycle, with pipeline architecture
- 1 Hardware Processor of considerable complexity - customized integrated circuits and a 1.3 Gbit integrated circuit memory
- 1 Vertically polarized antenna for RFI detection
- 1 Log Spiral polarized antenna for RFI detection
- 1 Commercial Spectrum Analyzer for RFI detection
- 1 Standard Large Disk 120 Mbyte, 7.5 Mbyte/sec R/W rate
- 1 Standard Tape Drive 10 Kbyte/sec R/W rate
- 1 Raster Scan Video Monitor with 3-D Analog Control
- 1 Microprocessor
- 1 Gbit total of Dynamic RAM

## System's Total Size

The OSD will fit in one medium-sized room. Existing prototypes of the four large capacity discs are surprisingly small, measuring slightly less than a standard size disc. It is estimated that the QWD will occupy an additional two racks, typically with one processor to a board, excluding the "Generalized Coherence by Rows," which will occupy several boards. Two additional racks will hold the Central Computer CPU and the Minicomputer/Display processor. This is shown in Figure 4.31.

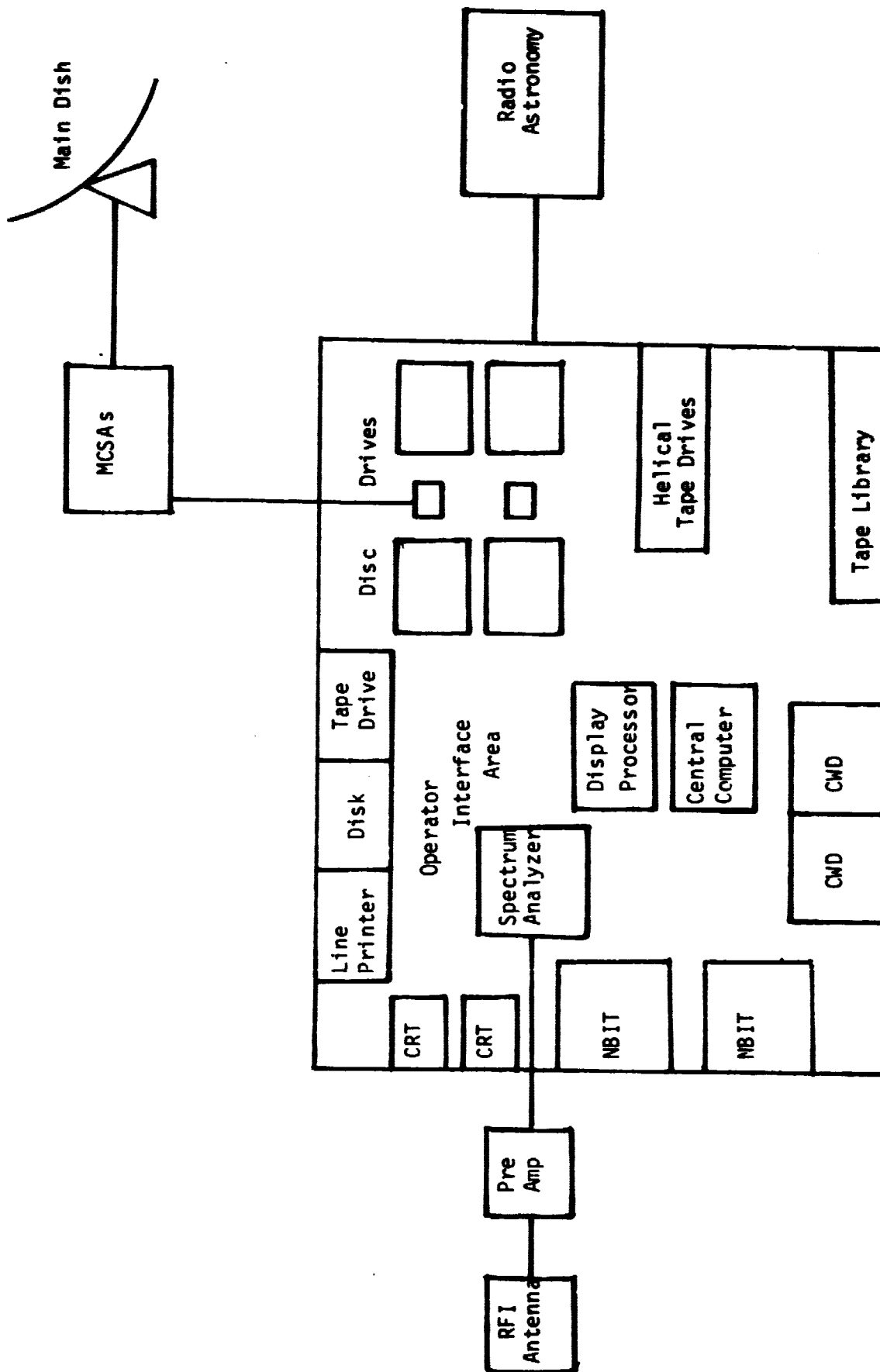


Figure 4.31 - The OSD Configuration

### System Cost

We list below the cost of the major components of the Oasis Signal Detector.

2 Tape Drives at 250K	\$500 K
4 Discs at 75K	300
2 Instant Replay Discs	150
1 I <sup>2</sup> S Terminal	200
All other RAM	120
Central Computer	10
	<hr/>
	\$1280 K

Since a general rule of thumb for digital design is that construction costs equal parts costs, the Oasis system would nominally be costed at three million (1979) dollars.



**Chapter 5**  
**Human Interface**

"Yes, I have a pair of eyes," replied Sam, "and that's just it. If they was a pair o' patent double million magnifying gas microscopes of hextra power, p'raps I might be able to see through a flight o' stairs and a deal door; but bein' only eyes, you see, my vision's limited."  
--Charles Dickens, Pickwick Papers

### 5.1 - Problem Statement

The human observer performs a number of key functions in the detection system both in the capacity of "pattern detector/recognizer" and as a decision maker who can redirect, verify, and supplement the automated phases of the system. The human interface system is an interactive link between the human observer (operator) and the rest of the system. It incorporates the observer in a variety of ways: as an alternative pattern detection algorithm, as one component of an algorithm which is mainly automated, as a decision-maker requesting information from the automated algorithms, and finally, as a trained scientist familiar with the astronomical context of the SETI program.

The way in which information is displayed often determines its actual usefulness. Effective decision-making about interesting ETI signals cannot occur unless that information is presented in a manner commensurate with an observer's perceptual and intellectual capabilities. The system uses an advanced image processor and multiple visual display screens as the primary modes of data presentation. An auditory display provides supplementary information.

### 5.2 - Using Human Information Processing Capabilities in SETI

One all-too-common approach is to automate every function that is technically possible and economically feasible. Whatever is left to do is relegated to the system operators by default. An alternative procedure,

and the one being advocated here, is to consider the human element as an alternative design component, complete with all the tradeoff characteristics one would identify in selecting one piece of hardware over another. The first question asked when considering the implementation of a search strategy or algorithm should not be how do we automate it, but, rather, should we automate it; if so, is total or semi-automation most appropriate?

The tradeoff for accepting some degree of inconsistency and low precision in performance specifications is that the human being as information processor is far and away the best model on the market, if it can be determined how to effectively use him. In developing the observer's role in the human-computer decision-making process, the following general attributes of human information processing are important.

1. Human performance often exhibits what has been termed "graceful degradation" under less than optimal conditions. That is, performance falls off gradually as conditions worsen, and significant (although not optimal) discrimination performance can be obtained even under very poor conditions. This can be contrasted with the discrete "go-no" paradigms characterizing most automated algorithms.

2. Human beings are conditional samplers of information. That is, the rate and type of information sampling can be continuously modified based on past experience and recent events. A human observer can frequently pick up clues to pattern phenomena by synthesizing events which, at the time they were perceived, were seemingly unrelated. Further, if he thinks he's "on to something", a human observer can then selectively attend to a narrowed range of signal stimuli, thus markedly enhancing sensitivity to those signals on a temporary basis.

3. People can process information on several levels. The "highest" level is conscious awareness. However, a tremendous amount of information

is coded and stored that is not necessarily accessible to conscious awareness; long-term storage is operationally infinite.

Even though an information display is "too large", presented "too rapidly" or lacks meaningfulness at the time it is viewed, it may be processed on a subconscious level to the extent necessary for the observer to refer to it in subsequent decision-making.

4. Certain attributes of a stimulus can be discerned even in the absence of sufficient information to permit identification. For example, color can frequently be perceived before form; motion before luminance. Thus, a human can scan large quantities of information very rapidly using certain stimulus dimensions (e.g., density) as a gross sort.

5. The trained human observer can develop and operate under very complex (by computer standards) decision criteria. For example, in a SETI observation, amplitude of extreme levels is obviously of initial interest regardless of other values (e.g., polarization). Below this extreme value, we might require additional parameters to attain given levels before the combination of information looks favorable enough to warrant detailed examination (e.g., type of polarization, consistency over time). Thus, an intricate system of multiple criteria which can be modified by experience can be developed by a skilled user. Although automated algorithms can be differentially weighted for various contingencies or combinations of variables, these weights must be derived prior to the search process and fixed arbitrarily at set values during the data analysis procedure. Humans can modify these weights "on line" under real time constraints.

6. Human algorithms can be altered by selection and training without the extensive hardware modification associated with some automated changes. It should also be noted that a human observer may be able to give information about why a currently used algorithm is not producing the desired

results; such qualitative judgements can be useful in interpreting more objective indices of system performance.

Straight lines and edges are very easily detected visually by humans. On the other hand, curved lines and dashed lines are less easily detected. In the present SETI context, it would therefore seem that a human observer would be adept at detecting a linearly drifting signal as shown as a display of the time-frequency output of the multi-channel spectrum analyzer (MCSA). On the other hand, a person may have more difficulty with pulses, and possibly greater problems detecting a wavy line embedded in a noise background. These predictions are indeed verified in our preliminary experiments (Appendix A5.1).

A more global characteristic of importance in human pattern detection is perceptual "symmetry." People note symmetry in visual patterns, prefer symmetry among components, and recall symmetrical patterns over time intervals more effectively than they do nonsymmetrical patterns (Garner, 1974; Chipman and Mendelson, 1979; Attneave, 1954; Szilagyi and Baird, 1977). Patterns symmetrical in time and/or space may be noted by the human observer over regions of the spectrum that are not compared by the automated algorithms described in Chapter 3. For instance, the human observer may note a similarity among clusters based on the N-bit representation of data blocks (where one pattern block = 200 x 50 data blocks). The simplest form of symmetry, of course, is recognition of the same pattern occurring at different time intervals but within the same region of the MCSA output space (say for different stars).

A further advantage of using the human observer as a detection device is his/her ability to deal with stimulus context. The letter A can be

written in a variety of type fonts, and yet the human can usually perceive (recognize) the characteristics on the page as an A. Machines have terrific difficulties with similar contextual problems, as attested to by the care with which numbers are stamped on bank checks to permit automatic pattern recognition! In the SETI program, the human observer will be harder to fool than the automated algorithms whose operations are more general than those of the "check processor". Additionally, the observer will be able to note the same pattern presented in a wide range of surrounding contexts.

In summary, the human as a detection algorithm will be effective at pulling out certain features of the MCSA output and results of the clustering scheme that will be missed by the automated algorithms. In this way, the observer both extends and supplements the automated aspects of the system. The role of the observer as a detection device is also advocated by Fender and Evans (1971) who review many of the advantages mentioned above, but in situations where the signal is known beforehand.

### Perceptual Sensitivity

It is anticipated that signals from extra-terrestrial civilizations, having traveled great distances through an uncertain electromagnetic environment, are of weak intensity as compared with background noise. The psychophysical models that deal with this type of situation as it applies to the human observer are part of the "theory of signal detectability" (TSD) (Baird and Noma, 1978; Green and Swetz, 1966; Swets, 1979).

According to statistical decision theory one can choose an optimum threshold if the values of all outcomes and the probabilities of signal and

noise present are known. This threshold is defined as

$$\beta = \frac{P_n}{P_s} \times \frac{N_n - Y_n}{Y_s - N_s} \quad (5-1)$$

where  $P_s$  = probability of signal plus noise being present

$P_n$  = probability of noise alone being present

$N$  = no response, "noise"

$Y$  = yes response, "signal"

$N_n$  = value of a correct rejection

$Y_n$  = cost of a false alarm

$Y_s$  = value of a correct detection

$N_s$  = cost of a miss

In the SETI context,  $P_n$  is much larger than  $P_s$ , thus suggesting that an observer would set a high  $\beta$ , and hence, would be disinclined to state a signal was ever present. But the second term of Equation (5-1) will tend to balance this bias. Looking first at the numerator, the positive value of a correct rejection,  $N_n$ , is small relative to the negative consequences of a false alarm  $Y_n$ . In the latter case the data must be archived, sorted and so forth--assuming  $Y_n$  is negative the numerator will be a positive number depending on the cost of archiving false alarms. This drives  $\beta$  up and reduces false alarms. However, the denominator more than offsets this tendency. The negative value of a miss  $N_s$ , though potentially large, merely represents our current level of ignorance, and is hopefully small relative to the value occurring from a hit  $Y_s$ . The latter term is obviously a subject for debate, but it is probably at least as high in respect to the negative outcome of a false alarm as the probability of noise is in respect to the negative outcome of a false alarm

as the probability of noise is in respect to the probability of a signal-- otherwise, we would not be initiating a search. Therefore we conclude that the far right hand term of Equation (4-1) is approximately the inverse of the middle term; thus leading to a moderate  $\beta$  of around 1.0. In other words an observer will not be fearful of suggesting the presence of potential signals, but will certainly be prevented from going overboard (claiming signals are everywhere) by the large value of  $P_n$  built up through past experience. Although the foregoing analysis is not intended to be rigorous, it does serve to highlight some of the very real factors an observer must consider in deciding whether or not a signal is present on the display at any instant in time.

#### Threshold Setting and False Alarm Rate

The preceding formula shows that the annoyance of false alarms is but one of four parameters that must be considered when establishing the thresholds of the algorithm. If we take into account all four trade-offs, we can determine the optimal threshold of the entire system and use this to set the thresholds of the individual processor alarms. For example, if the formula is used to determine that it is optimal to have at least one alarm sound from the system with 50% probability every 8000 seconds, on the average. To uniformly weight the individual thresholds  $T_{ij}$ ,  $i = (1, N)$ ,  $j = (1, M)$ , where  $N$  is the number of times an alarm may possibly sound in this period and  $M$  is the index of the particular alarm, we simply require that the probability that one will sound in 8000 seconds be  $P$ .  $P$  is then given by  $(1-P)^{NM} = .5$ , where we have assumed as an approximation that the alarms are independent.



This is the simplest arrangement. To weight different alarms, so that, for example, on the average the CWD alarm sounds more often than the PD alarm, we then require

$$\prod_{i,j}^{MN} (1 - P_{ij}) = .5$$

where  $P_{ij}$  is the weighted probability for the particular alarm.

The operator then, with the aid of a software routine, will select the mean time between alarms he desires, the relative probability for each alarm, and whether he would like the alarm to cause the central computer to automatically archive the entire observation or whether he will be contacted by the computer to pass the final decision (more usually the case).

#### Human Perception of Relative Magnitudes

Evaluation by humans of the processed receiver output data relies on the observer's ability to perceive the relative magnitudes of visual attributes such as brightness, length, position, and color, and auditory attributes such as loudness and pitch. When humans are given a reference level of some perceived quantity and told to assign that level a numerical value (say 10), they can give numbers to represent the relative perceived magnitude of other levels of that comparison quantity in respect to the reference level. Assuming these numbers are values along a ratio scale of perceived magnitude, the relationship (mapping) between physical and perceived magnitude can be described by a power function:

$$\psi = \lambda S^Y \quad (5-2)$$

where  $\psi$ , the perceived magnitude (numerical rating) is a power function of the physical intensity ( $S$ ). The exponent,  $\gamma$ , depends critically on the type of quantity being perceived. The multiplicative constant,  $\lambda$ , merely is a scale factor relating units of measure. The interesting parameter is the exponent. It is approximately .3 for both brightness and loudness, and closer to 1.0 for spatial position and length (Baird, 1970; Baird and Noma, 1978; Stevens, 1975). In other words, people are relatively accurate in their estimation of spatial position and length, whereas they grossly underestimate the relative intensity of lights and sounds. The practical importance of this result is that one gains little by having a CRT with  $2^8$  levels of brightness in a cell (pixel) since the human observer cannot resolve this many brightness levels. Similarly, we cannot anticipate high resolution of displays coded by auditory intensity.

Unlike quantitative attributes such as brightness, loudness and perceived length, sound frequency and hue are translated into qualitative perception dimensions. "Green" is not perceived to be twice "blue". According to Stevens (1975), these attributes do not follow the power law, and hence, we expect to use them in current system in cases where qualitative differences of low resolution are displayed.

Another psychophysical technique of relevance to our task is the method of absolute identification (Baird and Noma, 1978). Here a series of stimuli are presented sequentially and subjects must assign a unique label to each. By information theoretic measures, the channel capacity of the subject is determined for attributes such as light, sound and frequency. Surprisingly, subjects are only able to absolutely identify approximately 6 or 7 different magnitudes (without error) for unidimensional stimuli

(Miller, 1956). As the dimensionality of the stimulus pattern increases (say it varies in brightness, hue and size) the error-free level of identification goes up. In the present context, observers in the SETI program will be required to recognize radio frequency interference (RFI) at different points in time. We can expect such recognition capability to depend on the complexity of the RFI. The greater the number of dimensions one needs to define RFI, the greater the recall. For detection of signals themselves at low signal to noise ratios, we cannot expect the observer to be much help in recalling simple patterns over long time periods, or being able to discriminate a large number (certainly no greater than 10) noise displays. In brief, the human observer is limited in ability to recall simple targets over extended time periods, especially if intervening noise is present.

### 5.3 - Display Parameters Affecting Observer Performance

This section will discuss the major display variables associated with CRT's and their relationship to the human operator's search procedures. Many real-world tasks have been intensively studied (e.g., radar monitoring, piloting), and thus a considerable amount of information is known about the task dynamics.

#### Luminance

The luminance required for a CRT display is heavily dependent on contrast, ambient environment and the nature of the visual task. Gould (1968) cites 24 mL as a minimum figure, and recommends 50 mL for most applications.

Brightness (B) is the subjective impression of luminance (L). Under laboratory conditions the two are related by a power function (Equation (5-2)), where the exponent is typically between .25 and .50 for small, homogeneous fields (Marks and Stevens, 1966). This relationship does not hold for complex forms, however, particularly at higher luminance levels. The relationship appropriate for the SETI display will depend in large part on the scale of representation used (i.e., what each dot corresponds to), and on the mental set adopted by (or "trained into") the observer (i.e., how large, relative to the entire display, must a signal be?).

Another consideration for the SETI operator is the duration of the observing periods. Relatively high luminance levels may be optimized for short-term signal detections, but may also result in excessive glare and eye irritation over time.

In summary, the CRT display will be made highly adjustable in luminance and will be optimally reset for each graphical representation.

### Resolution

Assuming adequate luminance and contrast, a complex target image should subtend no less than 12 min. of arc (Grether, 1972). For "detail" resolution (e.g., one picture element) no less than 1.5 to 2.0 min. of arc is required (Bennett et al., 1967).

### Regeneration Rate and Viewing Distance

When the regeneration rate is too slow, the image appears to flicker. Angular size, luminance and phosphor persistence can all affect the

regeneration rate which is necessary to obtain a "critical flicker frequency" or CFF (the point at which no flicker is perceptible) for an observer (Kelly, 1965). Luminance and angular size alone can cause as much as 15 cps variation in the needed regeneration rate. When the individual differences in an observer's perceptual characteristics are added to the equation, it becomes obvious that there is a wide variability in the required regeneration rates for various applications.

Current image processors may use marginal regeneration rates because of the reciprocal relationship between the number of binary picture elements,  $N$ , and the regeneration rate,  $F$ , which is (in bits per seconds,  $I(\max)$  )

$$I(\max) = F \times N \quad (5-3)$$

Flicker can impact the observer's task in several ways. It can produce irritation and eye fatigue, particularly with the long viewing periods anticipated for SETI. It can also affect pattern perception indirectly by its influence over the viewer's distance from the screen. Thompson (1957) found that observers prefer viewing distances which are far enough from the screen such that small area flicker (e.g., that caused by the interlace between scan lines) is not resolvable. Thus, in addition to degrading pattern detection at relatively close distances to the screen, the presence of flicker may also determine the range of viewing distance adopted by the observer.

For the SETI application, an image processor with significantly less flicker than is typically found on CRT displays is indicated. A great deal of research is currently in progress to develop flicker-free image

processors (with adequate information display volumes) for the purpose of medical diagnosis. Special versions of image processors designed for medical applications are also marketed currently, and should be seriously considered in SETI. Some implications for human observers using the systems are given in Picket and Triggs (1975).

### Operator Training

The unique characteristics of the SETI application places considerable significance on the role of the system operator. Of particular interest for system designers is the potential long-term training effect. This could result in a specific observer becoming a much more effective signal detector over time. The SETI observer is not a typical "monitor". In a typical monitoring task where an operator monitors a CRT screen, and may conduct simple dialogues with the computer, a typical training curve would tend to asymptote after a few weeks. Any residual improvement as a function of practice for such tasks is insignificant for all practical purposes. However, in most monitoring tasks, "sufficient" or even optimal performance can be clearly defined when training begins; that is, if 100 signals are presented, 100 detections would equal perfect performance while perhaps 98-99 would equal "sufficient" performance. With the aid of guidelines and instructions, the learner goes through a series of successive approximations until criterion performance is reached.

For the SETI observer, no such feedback is possible. By the basic assumptions underlying the program, an ETI has never been detected. This implies that no direct measure of system performance (i.e., successful detection) will be possible for most cases. Therefore, the observers'

perceptions of their own performance may vary widely, and perhaps in an idiosyncratic manner. For example, one operator may decide to alter his search strategy because it has brought no results after 2 hours; another operator may follow a procedure for a significantly longer period before modifying or discarding it. One operator may feel that one strategy is more effective than another because it results in the selection of fewer total data regions for review, and thus is more sensitive; another operator may feel that this same approach results in too high a probability of missing a weak signal.

It should now be fairly clear that the training of SETI observers will be far from straightforward for the free mode of operation.

The training program is directed toward the following goals:

1. Observers are able to efficiently access and effectively use the display enhancement of the data. Data in various states of compression and analysis may be called up for viewing. Information from the different processors in the system will be assessed separately as well as interactively, such as might be required for a signal detection. Because each operator may put information from various sources together differently in a complex signal detection process, it will not always be clear what aspects of the information display are being utilized, and with what degree of effectiveness. By assessing the ability to use all information sources available, it can be determined if an observer can theoretically obtain the maximum information available for signal detection. Of course, this does not guarantee that all dimensions will be (or even should be) used in an observer's signal detection strategy.

2. Operators are encouraged to try alternate strategies. An analysis of search strategy archival records may suggest desirable patterns of search, which could be demonstrated to other observers.

3. An ongoing evaluation is conducted on veteran observers to determine if continuing experience enhances the ability to discriminate unusual (rare) patterns from more typical noise fields.

### Summary and Conclusions

The nature of the human-computer interaction in complex signal detection and recognition systems is perhaps the least developed aspect of such systems. The unique and virtually unexplored context of the SETI program makes research in this area a vital and potentially rewarding endeavor; not only for SETI, but for the areas of human-computer interaction, cognitive psychology and human factors as well.

The state-of-the-art in automatic pattern detection is far less advanced than scientists predicted five or ten years ago. The human operator will continue to represent the ceiling for system performance for the foreseeable future. As such, it is critically important that research be conducted which is aimed at deriving guidelines for observer task design, and for display-environment specification. The NATO Research Study Group on Pattern Recognition (Hodge, et al., 1979) cited the failure to integrate human factors considerations into automatic pattern detection systems as the major failure of systems design in this area.

### 5.4 - Human Factors Considerations

#### Design Philosophy

The human factors system analysis is required to ensure that the task procedures, environmental conditions and equipment designs are compatible with the psychological, physiological and performance characteristics of



the intended users. The design philosophy underlying this analysis can be summarized as follows:

1. Human performance for critical tasks must be, to the extent possible, designed into the system at the initial phases, not retrofitted after the fact.

2. Since humans are highly variable, there are virtually no configurations which are optimal for everyone. Systems can be designed to narrow the range of human variation through selection and training; the residual variation must be considered appropriately in the system design for human performance (the lower levels are subsumed under the advanced levels):

- A. Human Engineering: assuming adequate motivation, knowledge and skill, design the system to be operated in an optimal manner (easy, safe, reliable).
- B. Human factors: assuming adequate motivation, but not necessarily appropriate knowledge or skill, design the system to provide job aids and other support mechanisms to design compensate for human errors due to ignorance, task overload, environmental stress or improper decision-making.
- C. Behavioral Engineering: assuming nothing, design the system to predict the human failure modes including all errors (deliberate or intentional) due to neglect, sabotage or misuse, through one or more of the following three methods:
  - I. Prevention: make the undesired behavior less probable
  - II. Exclusion: make the undesired behavior impossible
  - III. Fail safe: assume that the undesired behavior will occur, and minimize the consequences to the system.

The level of design selected, and the method of implementation is a function of criticality and system resources. Table 5.2 outlines the major areas of analysis.

Again, the human factors function should be concerned with matching the tasks, procedures, equipment and environmental conditions comprising the Display Subsystem to the expected characteristics and capabilities of the system users in an "optimal" manner. (For further discussion of methodology, see Van Cott and Kinkade, 1972; or Chapanis, 1970). In this case, we assume that the range of skill, training, aptitude, motivation and performance manifested by system operators will be significant. Because operators may be drawn from a limited source (e.g., radio astronomers), the ability to modify these attributes through selection and training will be restricted.

As such, the limitations and predicted behaviors of the human operator become the "givens" around which the Operator Interface must be designed. The theoretical and expected capabilities of the operator serve as a ceiling on subsystem performance. From this ultimate value, the degree of degradation in operator performance caused by the operational conditions of the system (e.g., improper illumination, noise, task overload) must be predicted.

#### Summary of Some Key Human Factors Issues

1. The operator's task must be designed in a manner which will maintain active participation in the search process but which will not produce excessive false alarms, and will not result in overload-induced errors. This will require a detailed understanding of the intended

1. Physical Design and Work-space Layout

Ex.: Can the operator reach all controls and see all visual displays necessarily while performing sequential operations? Is the display-control ratio appropriate for the level of adjustment required on all image processors?

2. Workload and Task Requirements

Ex.: Is the workload sufficient to prevent boredom, fatigue and perceptual distortion but not so high as to create stress-induced errors?

3. Task and Procedure Design

Ex.: Is the operator receiving sufficient feedback to permit him to modify his performance (e.g., correct errors)? Is the appropriate sensory modality being used to display task-related information? Can one tell if the operator is performing as expected?

4. Emergency and Maintenance Procedures

5. Training and Selection Criteria

Ex.: What level of training is required for system operation? What training methods are appropriate? What individual characteristics should be considered desirable (or undesirable) for operator selection (e.g., spatial visualization ability)?

Table 5.2 - Major Categories of the Human Factors Functional Analysis

user groups. A description of relevant user characteristics should be completed prior to the finalization of system procedures.

2. The operator should be acquiring knowledge and experience by performing his role in the system, i.e., in addition to aiding the ongoing search process, the operator should be getting better at what he does over time (see "Training", in this chapter).
3. Tradeoffs between optimal display settings (e.g., luminance contrast) for "short-term" and "long-term" monitoring must be investigated since the operator's perceptual system will tend to adapt (selectively desensitize) over time.
4. The operator should have his performance monitored periodically to provide feedback and maintain attention allocation. This may involve the generation of test signals to see if the operator detects them. Test signals should be well above detection threshold to prevent a casual response set from developing.
5. Common failure modes must be anticipated and design compensated. For example, an observer who is absorbed in an active search process may forget to permanently record and save the data. System design must anticipate this type of event, and handle it by one of the three generic approaches (prevention, exculsion, or fail safe).
6. The role of system users is variable in the free mode so that skilled operators can fully exploit the system's capabilities. However, if naive or low-skilled operators are employed for fixed-mode tasks, these procedures should not require excessive interpolation or skilled judgments.

Table 5.3 outlines the key display parameters which are relevant to the human factors analysis while Table 5.4 summarizes the principal human factors considerations for the interactive search system. (The sequence in which items are found in Tables 5.3 and 5.4 does not reflect order of consideration or criticality).

1. Viewing distance
2. Angle of view
3. Image size and variability
4. Brightness and luminance
5. Contrast (range and proportion)
6. Visual field surrounding the display
7. Ambient environment
8. Resolution and image stability
9. Duration of a display pattern and interval length
10. Scale of dots (size of minimum informational unit)
11. Color
12. Single vs. multiple observers
13. Degree and type of visual cues (e.g., cursors)
14. Coding methods (e.g., shape, flicker, brightness)
15. Visual acuity required: level and type
16. Density of visual noise

Table 5.3 - Selected Visual Display Parameters Relevant to Human Factors Analysis

### 5.5 - General Structure of the Image Processor

The automatic signal detection algorithms alert the operator of a possible signal caused by RFI, random noise correlations or astronomical signal sources. This presents a massive data analysis and archiving

1. Display-Control compatibility
2. Population stereotypes; transfer of training
3. Rate of information input and requirements
4. Task loading
5. Reaction to predicted environmental stress levels
6. Training (skills required)
7. Operator background/experience
8. Speed vs. accuracy required (human reliability)
9. Information coding (e.g., shape, color, pattern)
10. Task design (e.g., sequency, coherence, dependency)
11. Communication requirements
12. Feedback and knowledge of results
13. Motivation
14. Work-space layout
15. Anthropometrics/biomechanics
16. Display mode selection and specification (for each task)
17. Maintenance/trouble
18. Design compensation for the disabled where possible
19. Work-rest cycles; circadian-biological rhythms
20. Emergency procedure/safety
21. Aesthetics (attitudes toward the work context)
22. Flexibility for future use
23. Public relations and system documentation (ability to demonstrate to naive persons)
24. Operator/system performance assessment
25. Behavioral engineering; resistance to intentional and accidental abuse

Table 5.4 - Representative Human Factors Considerations

problem and the operator's effectiveness in dealing with it depends upon the ability to quickly and efficiently view and manipulate subsections of the signal detector output and make correct decisions about the data presented. To further complicate the problem, the most interesting signals will probably appear at a very low S/N ratio, so sophisticated and versatile image enhancement techniques will be necessary to separate such a low-level signal from noise.

A high-speed digital image processor allows the operator to perform the various tasks associated with signal evaluation and system operation with high efficiency. In what follows, we describe the general characteristics required of a SETI image processor.

In response to a signal alarm, the operator requests that the subsection of signal detection output containing the signal be sent to the image processor's refresh memory. The refresh memory in a minimal system is two frames of static or dynamic random access memory (RAM), each providing three channels of 512 x 512 spatial resolution in 8 bit pixels (picture elements), and two 512 x 512 x 1 bit graphics overlay planes. This system allows for future expansion in resolution, number of frames and number of channels. As the need arises and as memory becomes cheaper, it will be possible to add more displays and/or enable the operator to scroll through a larger data base with fewer transfers from the host computer.

The graphics overlay planes store system status information for display on a black/white (B/W) status monitor to provide overlays such as graphs of statistical information or to designate subareas of the image for further processing.

A track ball or a digitizing tablet and stylus is used to position a cursor on the display. A trackball is a captive sphere that can be rolled

in any direction causing a corresponding movement of the cursor. When used in conjunction with function switches, individual pixels can be identified or an area can be outlined. The digitizing tablet and stylus can be used in a similar manner while additionally providing immediate hard copy on paper with an inking stylus. Scrolling through a zoomed image or an image larger than 512 x 512 is done with a trackball or joystick. The operator is able to scroll quickly by rows or columns through the entire refresh memory. Zoom is accomplished by positioning the cursor and specifying the zoom factor. Enlargement of at least 4:1 is done by pixel replication without altering refresh memory contents.

When data from both polarizations or two algorithms are to be viewed simultaneously, a split-screen view is used. This is accomplished by selective loading of refresh memory by the host computer or by reading subsets of two different refresh memories. A feedback loop allows the processed image to be written back into refresh memory for iterative operations and spatial convolutions. The operator can perform sophisticated image enhancement functions in evaluation potential signals without involving the host computer.

A built-in microprocessor handles all internal control functions as well as some processing operations, thus off-loading the host computer of most image-oriented tasks. Microprocessor programs will be down-loaded from a library on the host computer. The microprocessor performs diagnostics and aids in locating problems within the image processing system.



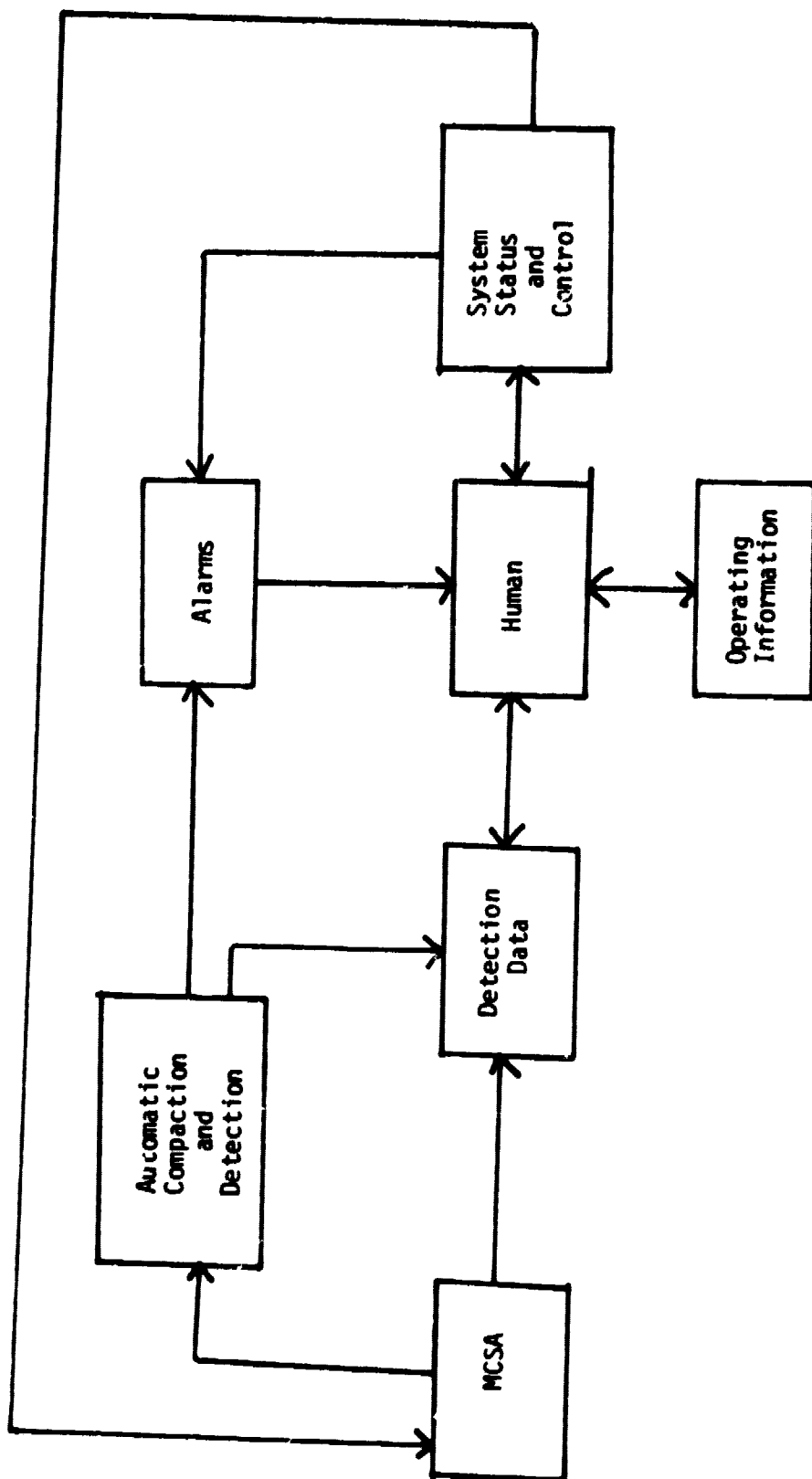


Figure 5.1

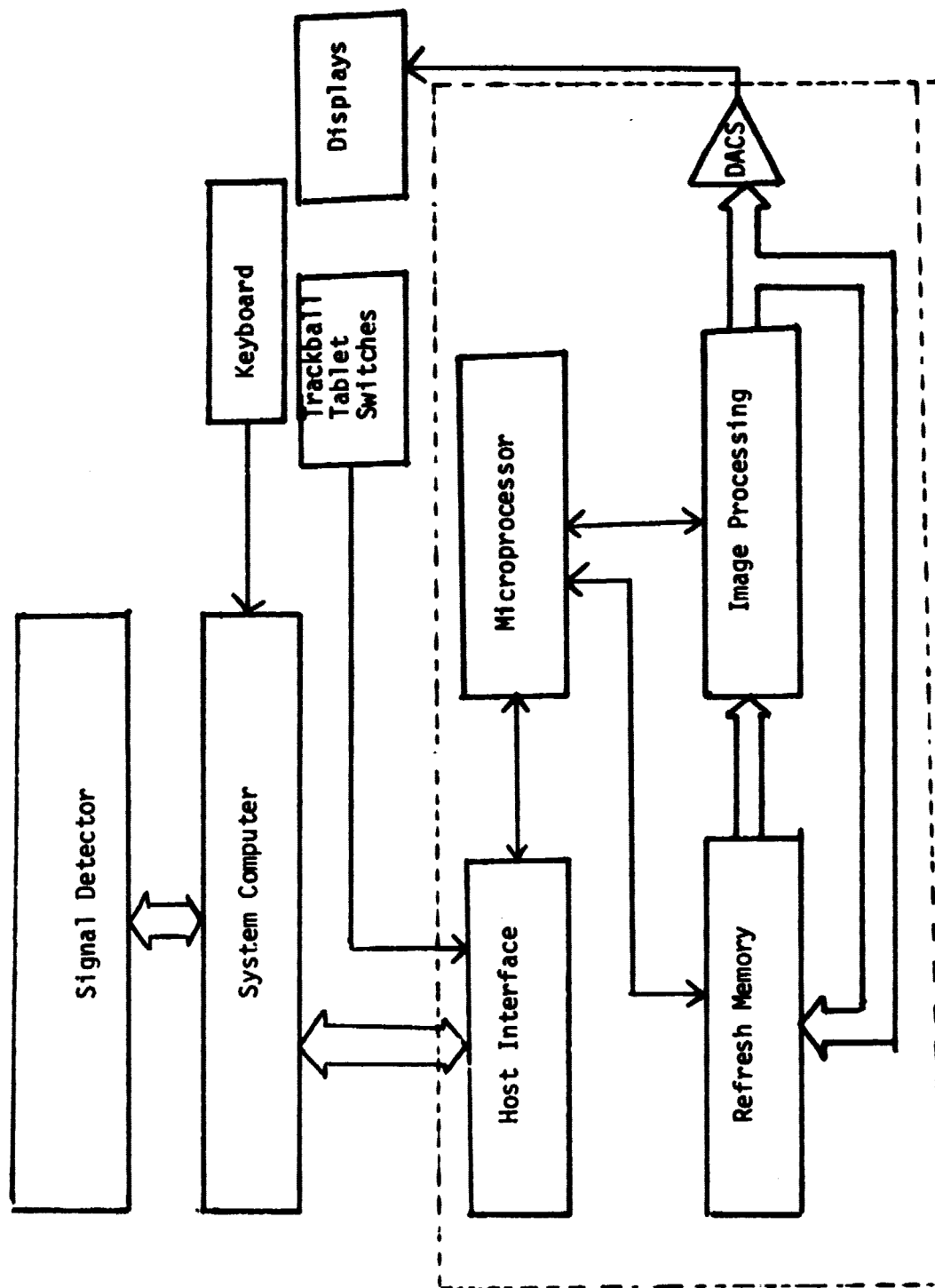


Image Processor Conceptual Diagram

Figure 5.2

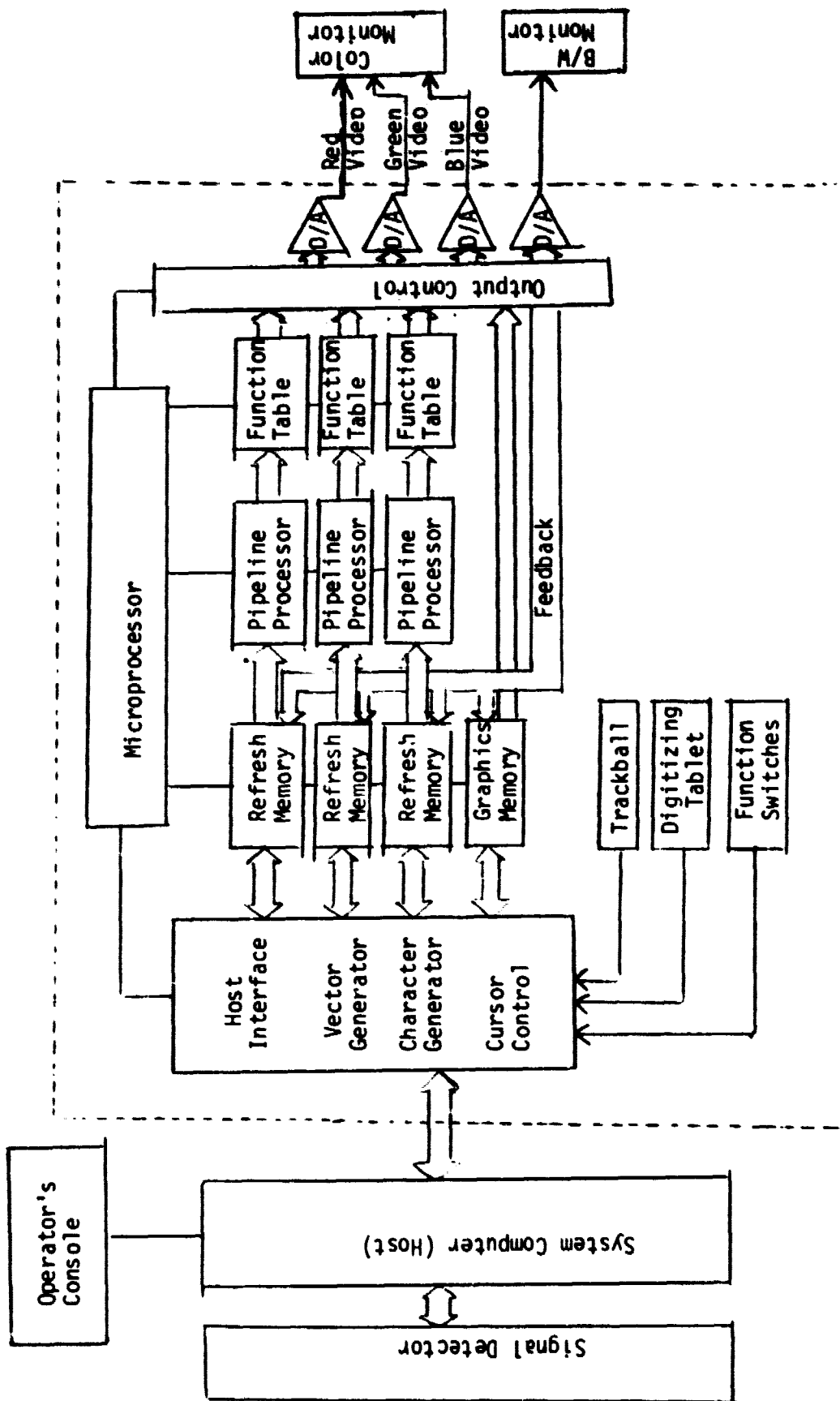


Image Processor

Figure 5.3

## 5.6 - System Overview

The displays subsystem must perform four types of functions: system status, operating information, alarms, and detection.

System status is the operator's bookkeeper. It keeps track of all information concerning system faults and operating parameters (target star, gain, bandwidth, etc.). Further, it provides for input regarding these parameters by the operator. Operating information includes a library of procedures and catalogues readily accessible from airplane schedules to "How to Replace a 16K RAM Board (In three easy steps)". Alarms signal the human that the machine requires some attention, whether it be to check a possible ETI detection, to see why the MCSA malfunctioned, or to perform maintenance operations (e.g., put more paper in the line printer).

Detection is really what the entire system, and indeed SETI itself, is all about. This includes everything the operator uses to either search for ETI signals or to cull out RFI from automatically "detected" signals. The display parameters are readily changeable by the operator. Complex analyses, such as 2-D Fourier or Walsh transforms, that are too expensive, time consuming, or experimental to be done on the entire output can be done quite easily on a  $512^2$  display matrix. Such support systems could aid the operator significantly as the SETI program advances into more sophisticated analyses of the radio spectrum.

There are two major modes of operator observation, "fixed" and "free". By fixed mode we mean the operator responds to the results of the automated algorithms and system status test. In this mode he/she monitors equipment for proper functioning, and when a detection occurs (alarm from automated algorithms) it is analyzed, and appropriate action is initiated. In the

"free mode" the operator acts with greater freedom concerning search strategies. For example, one may monitor a selected region of the MCSA output displayed on the color CRT, listen to a selected region of the spectrum, or call up statistical analyses associated with specific tests in the N-bit number.

### Fixed Mode

Although the flexibility of the human observer in formulating, modifying, and executing search strategies is a major resource for SETI, less flexibility is required for the fixed mode. Here the operator performs a relatively fixed set of procedures, especially in response to an alarm. Failure to attain high levels of standardized performance would result in a non-systematic data base which would hinder the cumulative and longitudinal nature of the SETI program. There are two phases of the fixed mode: normal and examination.

Normal State - Between alarms the operator is provided with both visual and auditory displays that give a selected sample of all incoming signals, as well as status information on the observations and equipment operation. The color CRT displays a subsection of the MCSA output spectrum. Each spectrum is displayed as a horizontal line, with the intensity of a pixel proportional to the average signal power at that frequency. Each second, a new spectrum is added beneath the existing ones. In this way up to 512 seconds of output may be observed simultaneously. Since there will be two MCSA spectra (one for each polarization), the display will alternate between them at the discretion of the operator.

The auditory portion of the system serves as a final step in checking a possible signal of interest which has been flagged by an algorithm or by

the operator. A search of the entire 8 MHz band each second is not undertaken due to the complete inability of the operator to distinguish differences at such a rapid rate. Since an alarm will not yield information as to which MCSA data caused the flag, the computer-controlled tunable system checks the real part of the input for each MCSA.

The auditory system taps into the data flow at an intermediate frequency (IF) before the MCSA. Using an appropriate local oscillator, mixer, and 26-34 MHz bandpass filter as shown in Figure 5.4, the IF signal is heterodyned down before being received and converted to the auditory range of 0-8 KHz. As alarms are sounded for possible ETI's, the computer controlled system places the indicated algorithm frequency in the middle of the audio band. For algorithms using 40 Hz by 20 sec. blocks, the center block frequency is the one used.

Based on cost, the pre-MCSA data are not stored. Since the auditory system is in real time, the operator hears regions which are down stream "n" seconds from when the alarm was triggered. Therefore, the signal may have drifted off the alarm frequency by the time the operator "hears" it. The number of seconds "n" may vary from a few seconds to 1000 seconds. For the worst case of  $n = 1000$  seconds, and a drift rate of 1 Hz/sec., the signal would drift only 1 KHz and would still lie within the 8 KHz band. To check for more rapidly drifting signals, and to allow for error, the receiver tunes to the indicated frequency and to adjacent 8 KHz bands. The auditory display correlates real time with the past, and yields a check for the continuing presence of the detected signal. In alarms which originate from the cluster analysis ( $n = 1000$ ), the operator may elect to hit the override button and hold the antenna in the location of interest.

# Auditory System

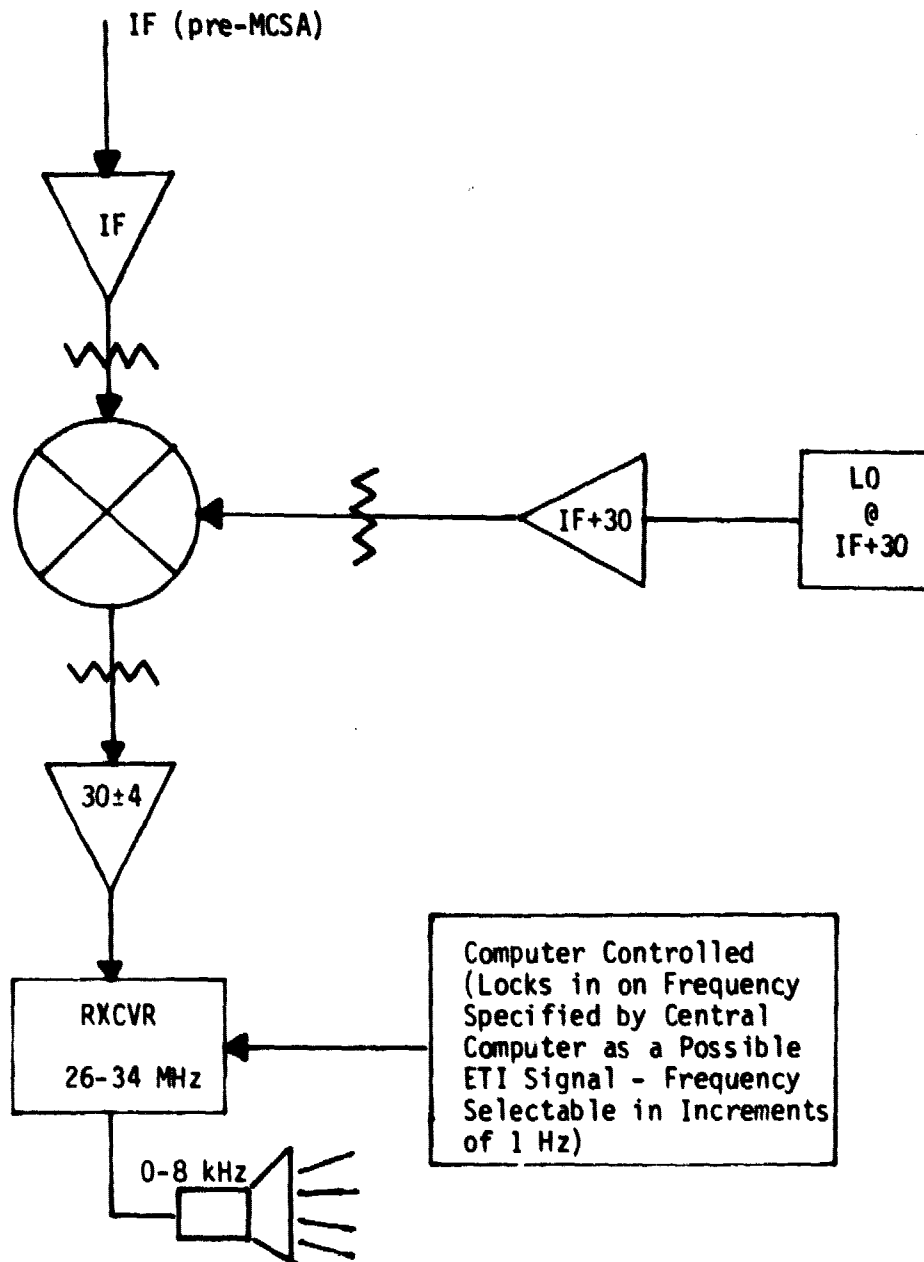


Figure 5.4

A second black and white CRT continuously displays a variety of system-status data. Some of this information is shown only upon request. This information is also recorded in permanent archives whenever an alarm is signaled by either the operator or the automated algorithms. Typical status parameters are:

1. Observing Parameters

- a. Telescope pointing direction
- b. Target star name, distance, type, etc.
- c. Sky map of the area around the antenna beam (about 10 beam-widths) including all known astronomical objects, at all wavelengths
- d. Date, time, polarization
- e. Operator's name

2. Receiver Parameters

- a. System gain and bandpass shape
- b. Self-diagnosis information. Tests must be end-to-end.
- c. MCS input and output bandwidths.
- d. Status of any selectable option

3. RFI Parameters

- a. Radio and TV catalog and schedules
- b. Earth satellite catalog and schedules
- c. Airline schedules and flight paths

4. Environmental Parameters

- a. Wind, precipitation and lightning indicators
- b. Outdoor TV camera, steerable and zoomable, to allow watching the antenna and its surroundings for unusual conditions that



could cause false alarms or malfunctions (snow drifts, flocks of birds or insects, trespassers, etc.)

Examination State - When an alarm occurs from the automated algorithms the operator carries out a predetermined set of procedures, somewhat different for different types of alarm. For example, the drifting carrier detector (CWD) triggers an alarm to the operator whenever the integrated coherent power along a drift line exceeds the present threshold. The operator first requests a graphical output showing the location of all data in the vicinity of that which caused the alarm. At this point, and at each succeeding one in the decision process, the operator may enter a coded value at the CRT controls regarding the nature of the alarm. There are four possibilities: noise (N), radio frequency interference (RFI), useful astrophysical information (AI), potential signal from extraterrestrial intelligence (ETI). Selection of option 3 or 4 saves the region of interest in a permanent record and automatically records all current status parameters. Option 1 or 2 terminates consideration of the alarm. If the ETI option is selected, the operator must decide whether or not the telescope should be kept on the star beyond the 1000 seconds normally allotted.

Before an option is chosen the operator may request further data concerning the alarm. Specifically, through instant replay the original MCSA output can be brought onto the color CRT for more detailed examination. In this case, the output is coded in a frequency x time matrix, with amplitude represented by pixel brightness. By the use of split screen techniques, both MCSA outputs of the same region are shown. Through the use of the CRT, the operator can zoom into a subregion, produce a histogram of

the intensities, and have the appropriate vectors color-coded and displayed. The purpose of all such tests is simply to speed and aid the final decision-making process with respect to the four options (N, RFI, AI, ETI). This selection can be made at any time during a search with the consequent procedures outlined above.

The operator also receives auditory information about the current state of affairs in the region of interest. As soon as an alarm is signalled to the operator, the communications receiver driving the auditory display is automatically tuned to the signal frequency (origin of the vector) and its bandwidth adjusted to encompass a 8 KHz region centered on the origin. Since the auditory pickup is in real time, the operator is able to ascertain whether the signal is still present in the region of the spectrum delimited by the instant visual replay. This sort of data will be helpful in deciding whether further observing time should be devoted to that star region. In addition, the operator enters any special comments about the circumstances of the alarm via the hard-copy terminal.

A second type of alarm is one triggered by the clustering algorithm based on grouping similar N-bit patterns. When the operator receives such an alarm, the instant replay tape is searched for the pattern block of interest, as well as the two adjacent pattern blocks (adjacent in the frequency spectrum). A pattern block is a 200 x 50 matrix of data boxes. Each data box is a 40 Hz x 20 second region of the MCSA output. The pattern block groups 200 (overlapping by 8 Hz) frequency blocks of 40 Hz each by 50 time blocks of 20 seconds each.

The initial display of these data is of the critical pattern block in the center of the color CRT flanked and surrounded by the eight adjacent

pattern blocks (appropriately translated to prevent display of redundant data, and clipped on the left and right sides to fit on 512 x 512 screen). Since a pattern block is based on data from both MCSA's, only a single output is shown on the screen. The location of the N-bit numbers of interest are shown as individual pixels of a constant brightness against a dark background. Details of the statistical tests comprising the N-bit number are printed onto the hard-copy terminal. An example of this display in tabular form is given in Table 5.1. At this junction, the operator proceeds as previously described: decides which category is appropriate to define the alarm, or request further information.

Tests	Status	Significance
1. complex coherence	1	$>\alpha_1$
2. broadband pulse	1	$>\alpha_2$
3. degree of polarization	1	$>\alpha_3$
4. goodness-of-fit	1.	.01
5. ANOVA rows	0	.20
6. ANOVA columns	1	.001
7. ANOVA interaction	0	.20
8. total power	1.	$>\alpha_4$
9. 8-Hz pulse	0	$<\alpha_5$

$\alpha_i$  = threshold parameters to be given numerical values (for different levels)

Table 5.1 - Sample Output to Represent Results of N-bit Pattern

If a decision cannot be reached immediately, the operator requests that closely related bit patterns be displayed simultaneously within the pattern blocks. In this case, the alternative bit patterns are coded with different

colors. Finally, while still at the level of the pattern block, the criterion or statistical level at which an individual test of the N-bit pattern was passed is displayed as a brightness value (higher brightness represents higher significance).

If more detailed results are still necessary before a decision can be reached with any confidence, the operator specifies the relevant region of space within the pattern block by cursor control, and the original data boxes (40 x 20) in that region are displayed. By split screen techniques, both MCSA outputs are shown side by side. Simultaneously, the auditory display is automatically turned on and moved to the same region of the current frequency spectrum. (This may be turned off, lowered in volume, and so forth, at the discretion of the operator). From this vantage point, the zoom, histogram and color coding options of the CRT are used to highlight potential signals and eventually aid in the operator's decision concerning the underlying cause of the alarm.

#### Free Mode

In addition to responsibilities associated with the fixed-mode displays, the operator carries out analyses of possible signals in a free mode. Indeed, most of the time an observer will be viewing or listening to a noise field that does not contain a signal. In the free mode, the observer may think a signal is present, based on the existence of an unusual pattern perceived on occasion within the display. In the SETI situation, such "thoughts" may in fact reflect bona fide signals.

The real advantage of the free mode comes through the flexibility of the human being as a pattern detector. The inherent limitation of the automated

algorithms is that features of potential patterns must be specified beforehand. The operator works under no such constraints. He is given virtual freedom to explore any means which will lead him to make a decision. These include

3 archives

- Entire Observation Archive
- Special Interest Archive of previous observations
- Pattern Block Display

RFI detection system memory

Radio Astronomy System, Radio Astronomer

Software routines emulating system algorithms

High level software routines of his own design

His own memory, insight and intuition

In conclusion, the free mode allows the detection system to operate at its best, combining the broad and diligent scrutiny of the machine, with the wisdom, judgment, and inventiveness of the human.

## Comparison of the Fixed and Free Modes of Operation

Characteristic	Fixed Mode	Free Mode
1. Nature of Interaction	<ul style="list-style-type: none"> <li>- Directive</li> <li>- Evaluative</li> </ul>	<ul style="list-style-type: none"> <li>- Informative</li> <li>- Responsive</li> </ul>
2. Information Needed	<ul style="list-style-type: none"> <li>- Search Process Steps</li> <li>- Feedback; Errors</li> <li>- Time Elapsed</li> <li>- Time Remaining</li> </ul>	<ul style="list-style-type: none"> <li>- Search Options</li> <li>- Archival Records</li> <li>- Current Search Record</li> <li>- Time Elapsed</li> <li>- Time Remaining</li> </ul>
3. Operator Task Load	<ul style="list-style-type: none"> <li>- Low</li> </ul>	<ul style="list-style-type: none"> <li>- Low to High</li> </ul>
4. Operator Task Sequence	<ul style="list-style-type: none"> <li>- Fixed</li> </ul>	<ul style="list-style-type: none"> <li>- Variable</li> </ul>
5. Operator Task Mode	<ul style="list-style-type: none"> <li>- Forced Paced</li> </ul>	<ul style="list-style-type: none"> <li>- Self-Paced</li> </ul>

## References

1. Attneave, F. Some informational aspects of visual perception. Psychol. Rev., 61, 1954, 183-193.
2. Attneave, F. Symmetry, information and memory for patterns. Amer. J. Psychol., 66, 1955, 209-222.
3. Baird, J.C. Psychophysical Analysis of Visual Space, Pergamon, Oxford, 1970.
4. Baird, J.C.; and Noma, E. Fundamentals of Scaling and Psychophysics.
5. Bennett, C.A., Winterstein, S.H., and Kent, R.E., Image Quality and Target Recognition, Human Factors, 9, 5. (1967).
6. Pieberman, L.M.; Fallacy and Fact of Sampled Image Displays, Human Factors, 16, 286-299. (1974).
7. Bieberman, L.M. (Ed.) Perception of Displayed Information; Plenum Press, London. (1973).
8. Chapanis, A. Human Factors in System Engineering, In: Systems Psychology, DeGreene, K. (Ed.), McGraw-Hill, New York, (1970).
9. Chipman, S.F.; and Mendelson, M.J. Influence of six types of visual structure on complexity judgements in children and adults. J. Exp. Psychol.: Human Perception and Performance, 5, 1979, 365-378.
10. Fenker, R.M.; and Evans, S.H. A model for optimizing the effectiveness of man machine decision making in a pattern recognition system. U.S. Army technical memorandum. 8-71, 7p, 1971.
11. Fuld, K. A sensitization effect with rectilinear stimuli. Vision Research, 18, 1978, 1045-1051.
12. Garner, W.R. The Processing of Information and Structure. Erlbaum Assoc., N.Y., 1974.
13. Gould, John D., Visual Factors in the design of computer controlled CRT displays Human Factors, Vol. 10, Number 4, 359-376. (1968)
14. Green, D.M.; and Swets, J.A. Signal Detection Theory and Psychophysics. Wiley, N.Y., 1966.
15. Grether, W., Visual Presentation of Information, Human Engineering Guide to Equipment Design, H. Van Cott and R. Kinkade (Eds), American Institute for Research, Washington D.C. (1972).
16. Hairfield, H.W. Night and All-Weather Target Acquisition: State of the Art Review, Part III: Television and Low-Light-Level Television Systems, Boeing Company Report D162-10116-3., (1970).
17. Hodge, D.C. et. al. NATO Research Study Group on Pattern Recognition: Final Report. U.S. Army Human Engineering Laboratory., Aberdeen Proving Ground, Maryland. (1979).

## References(cont.)

18. Hubel, D.H.; and Wiesel, T.N. Receptive fields and functional architecture of monkey striate cortex. J. Physiol., Lond., 195, 1968, 215-243.
19. Kelly, D.H. Flicker Thresholds, In P.W. Nye (Ed.) Information Processing in Sight Sensory Systems, C.I.T. Press (1965)
20. Lindsay, P.H.; and Norman, D.A. Human Information Processing. 2nd ed., Academic, N.Y., 1975.
21. Miller, G.A. The magical number seven, plus or minus two: Some limits on capacity for processing information. Psychol. Rev., 63, 1956, 81-97.
22. Pickett, R.M. and Triggs, T.V., Human Factors in Health Care, D.C. Heath and Co., Lexington. (1975).
23. Reed, S.K. Psychological Processes in Pattern Recognition. Academic, N.Y., 1973.
24. Rosell, F.A. and Willson, R.H., Recent Psychophysical Experiments and the Display Signal-to-Noise Ratio Concept, In: The Perception of Displayed Information, L.M. Bieberman, L.,. (Ed.), Plenum Press, London. (1973).
25. Selfridge, O. Pandemonium: A paradigm for learning. In Symposium on the Mechanization of Thought Processes. H.M. Stationary Office, London, 1959.
26. Snyder, H.L., Image Quality and Observer Performance, The Perception of Displayed Information, L.M. Bieberman (Ed.), Plenum Press, London. (1973).
27. Snyder, H.L. et. al. Low-Light-Level TV Viewfinder Simulation Program. Phase A: State-of-the-Art Review, USAF Report AFAL-TR-67-293. (1967).
28. Stevens, S.S. Psychophysics: Introduction to its Perceptual, Neural and Social Prospects. Wiley, N.Y., 1975.
29. Swets, J.A. ROC analysis applied to the evaluation of medical imaging techniques. Investigative Radiology, 14, 1979, 109-121.
30. Swets, J.A.; Green, D.M.; Getty, D.J.; and Swets, J.B. Signal detection and identification at successive stages of observation. Percept & Psychophys., 1978, 23, 275 - 289.
31. Szilagyi, P.G.; and Baird, J.C. A quantitative approach to the study of visual symmetry. Percept. & Psychophys., 22, 1977, 287 - 292.
32. Thompson, F.T. Television Line Structure Suppression, Journal of the Society of Motion Picture and Television Engineers, No. 66, 602-606. (1957).
33. Van Cott, H., and Kinkade, R., (Eds), Human Engineering Guide to Equipment Design, American Institute for Research, Washington D.C. (1972)
34. Zusne, L.: Visual Perception of Form, Academic, N.Y., 1970



**Chapter 6**  
**Radio Astronomy Applications - Molecular Spectroscopy**

Some things you miss because they're so tiny you overlook them. But some things you don't see because they're so huge. We were both looking at the same thing, seeing the same thing, talking about the same thing, thinking about the same thing, except he was looking, seeing, talking and thinking from a completely different dimension.

--Robert Pirsig, Zen and the Art of Motorcycle Maintenance

As of the time of this writing, there are nearly 50 molecular species, ions, and radicals that have been found in the interstellar medium by radio astronomers. Observations of the lines from these molecules provide important astrophysical information about the sources in which they are found. As well as being discovered in gas and dust clouds in the interstellar medium, molecules have also been observed in circumstellar shells. There is a constant recycling of matter between stars and interstellar space, with new stars being born from the gas and dust and old and dying stars releasing matter back into the interstellar medium, molecules have also been observed in circumstellar shells. There is a constant recycling of matter between stars and interstellar space, with new stars being born from the gas and dust and old and dying stars releasing matter back into the interstellar medium. Currently we estimate that about 90% of the mass is in stars and the rest in gas and dust that occupies the volume between the stars.

From many of the compounds detected, radio astronomers are able to observe several different transitions or lines. From the intensities and widths, physical conditions such as temperature and density may be probed. Calculations can be made to estimate the column density of a specific molecule and hence provide relative abundance information. Abundances are better determined when several lines are available from different energy levels so that the population statistics in various energy levels can be

estimated. The presence of a molecule together with its abundance provides fundamental data for interstellar chemistry, that is the reactions and rates responsible for production and destruction of the molecules.

For each molecule present there is also information that is automatically available for velocity along the line of sight to the observed region from observations of the Doppler shift. A grid of observations in a molecular line then provides us a view of the kinematics of the region. In addition to the nearly 50 different molecular species found to date, which so far are composed only of atoms of H, C, N, O, Si, and S, there are also innumerable isotopes discovered. These add to the list of atoms the following: D,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ,  $^{17}\text{O}$ ,  $^{18}\text{O}$ ,  $^{29}\text{Si}$ ,  $^{30}\text{Si}$ ,  $^{33}\text{S}$ , and  $^{34}\text{S}$ . By observing line ratios of different isotopic species it is possible to obtain isotopic abundances when the cloud is optically thin. In turn, the nucleosynthetic history of our galaxy is highly dependent on isotope ratios such as  $^{12}\text{C}/^{13}\text{C}$  and  $^{17}\text{O}/^{16}\text{O}$  (see for example Audouze, J., and Tinsley, B.M., 1976 Annual Review Astronomy and Astrophysics).

The characteristics of the MCSA/SD make it uniquely useful for some radio astronomy applications. In other instances where it is not necessarily unique it will be a useful back end for line work. Since we feel the primary use of the instrument will be in conjunction with SETI in the water hole, we emphasize work that could be done in this frequency range. However, the MCSA/SD has much more general capability so we also discuss projects of interest that are not restricted to the water hole. In addition the MCSA being built by JPL is different from that being built for NASA-Ames, and different problems lend themselves better to one or the other.

Table 6.1 contains a list of molecular lines in the water hole observed in one or more astronomical sources. We list in Table 6.2 other lines in the

water hole that have not yet been detected astronomically but are of astro-physical interest since other lines from the species or an isotope have already been observed. The ultra high resolution capability will in most instances not be necessary for work on these lines. Thermal widths in typical cold clouds are larger than about 1 KHz while the actual widths in even the narrowest lines will be larger because of turbulent broadening. In a search for new lines the JPL MCSA will be extremely powerful. With a total bandpass of 300 MHz and  $10^6$  channels, the spectral resolution will be 300 Hz. This provides sufficient resolution while covering a bandpass as large as the waterhole in one setting. In this regard an instantaneous bandpass of 500 MHz would be even more desirable. The detection of new molecules will provide additional pieces in the puzzle of the young science of astro-chemistry. Radio astronomers are on the verge of detecting amino acids in space. Glycine, the simplest amino acid, has 10 atoms and a molecular weight of 75. Molecules as complicated as 11 atoms and a molecular weight of 126 have already been found by radio astronomy spectroscopy. In addition, all of the chemical bondings that are present in glycine have been seen in at least one of the molecules found already.

Current OH surveys in our galaxy have been limited to  $\pm 2^\circ$  of the galactic plane and have been unable to cover high velocities because of limited bandwidth. With the JPL MCSA all four OH lines could be covered simultaneously (as well as lines of the isotopes) perhaps making it worthwhile to undertake a major new OH-survey, that would also pick up the high velocity OH. As long as we know the positions of OH to within  $1' - 2'$  very useful polarization work could be done on the OH masers. These new observations could provide the necessary resurgence for theoretical work on classification schemes and pumping mechanisms.

An interesting experiment with the NASA-Ames MCSA will be to attempt to find Zeeman splitting in OH. Crutcher and Heiles (private communication) feel they have a possible first detection of this for OH from a dark cloud in Taurus seen in absorption against the source 3C133. They observed both right hand circular (RHC) and left hand circular (LHC) polarization with the Arecibo telescope. With two MCSA's providing RHC and LHC polarization, we would subtract one spectrum from the other and look for the slightly shifted lines which would manifest themselves by a sideways S profile. The height of the ripple is proportional to the magnetic field, and the frequency difference between maximum and minimum is equal to the full width at half maximum, FWHM, of the line. The narrowest lines found so far in 18-cm OH are about 0.5 Km/s or 3 KHz wide. This is the case for the dust cloud L134N. This is considerably greater than the thermal width which for a 10K cloud is about 0.1 Km/s. Chances of successful detection will be enhanced when clouds are discovered that are both colder and less turbulent or have a larger magnetic field. The Zeeman splitting is 3.27 Hz/ $\mu$ G for the 1665 MHz line and 1.96 Hz/ $\mu$ G for the 1667 MHz line (see e.g., B.E. Turner and G. Vershuur ApJ 162, 341, 1970). Zeeman splitting in the 21-cm HI line has been measured in a number of sources since the first detections by Vershuur (ApJ 156, 861, 1969 and Nature 223, 140, 1969). The splitting is 2.8 Hz/ $\mu$ G. Further work could profitably be done on HI as well as on other molecules with an unpaired electron spin such as CN, CH, and SO that would have comparable Zeeman splitting.

The spectral resolution provided by the MCSA/SD will easily resolve nuclear quadrupole hyperfine splitting. This has already been done for formamide but not yet for  $^{13}\text{C}$  formamide (see Table 6.1). The latter would require

a more sensitive system than available to the French group that measured one component of  $\text{NH}_2^{13}\text{COH}$ . Even when the rotational energy levels are not populated according to local thermodynamic equilibrium LTE as is the case in formamide, the hyperfine components appear to have relative intensities conforming to LTE. This would be interesting to corroborate and further check in other molecules. The MCSA/SD system would be particularly useful for more closely spaced hyperfine lines such as the spin-rotation lines of the  $1_{10} \rightarrow 1_{11}$  transition at  $\text{H}_2\text{CO}$  at 6-cm. The  $F_u \rightarrow F_L$  lines  $0 \rightarrow 1$  and  $2 \rightarrow 2$  are separated by about 1 KHz. Astronomically, they appear to be in LTE.

As mentioned above the nucleosynthetic history of our galaxy depends heavily on ratios of line intensities of various isotopic species. According to A.A. Penzias (IAU Symposium No. 87 Interstellar Molecules meeting 8/1979 to be published) the interstellar  $^{13}\text{C}/^{12}\text{C}$  abundances ratio is about 150% of the terrestrial value of 1/89 throughout the plane of our galaxy and about twice the terrestrial value near the galactic center.  $^{12}\text{C}$  is a primary element formed directly from H and He inside a star while  $^{13}\text{C}$  is considered a secondary element produced from primary elements in second and later generation stars by the S-process. Table 6.1 and 6.2 show that isotope experiments could be done with lines in the water hole.

The MCSA coupled with the Signal Detector would provide a valuable tool for radio frequency pulsar investigations. This apparatus makes possible the high resolution power spectrum monitoring of pulsar pulses, as well as a frequency domain record of the type and degree of polarization (stokes parameters) associated with a single pulsed period. The following experiments are suggested as worthwhile:

1. Measuring the bandwidth of micropulses and nanopulses within a single pulse.
2. Using Power spectrum, or autocorrelation functions derived from the Power spectrum, to determine the relative contributions of sub-pulses, micropulses, and nanopulses to the total integrated single pulse.
3. Exploring the possibility of extraterrestrial transmitters being located near a pulsar, using the pulsar as the signpost advertising a beacon signal.

### Radio Recombination Lines

Since radio recombination lines are very wide there is no need to observe them with the MCSA/SD having 8 MHz bandpass and  $8 \times 10^6$  channels. On the other hand an instantaneous bandpass of about 300 MHz with  $10^6$  channels would provide a unique combination for recombination line experiments. For example, the entire water hole could be covered with a single bandpass that would provide simultaneous observations of the following hydrogen, helium, and carbon recombination lines:

$\alpha$ - lines	156 - 166
$\beta$ - lines	196 - 209
$\gamma$ - lines	224 - 238
$\sigma$ - lines	247 - 282
$\epsilon$ - lines	265 - 282

Rest frequencies for these lines can be found in (A.E. Lilley and P. Palmer ApJ Supplement 144, 16:143, 1968). Such observations would provide better data in the sense of good internal calibration because of the simultaneity of

the observations. A new level of accuracy may then be brought to bear on such familiar problems in gaseous nebulae as:

- 1) local thermodynamic equilibrium (LTE) and departures therefrom.
- 2) electron temperatures
- 3) line broadening - Stark effect
- 4) gas dynamics - within an individual object as well as an aid in galactic structure determination.
- 5) relative abundance of helium to hydrogen with its cosmological significance.
- 6) carbon line studies to determine regions of emissions compared to hydrogen and helium, and relative abundances.

For additional commentary on the above avenues of research see for example

R.L. Brown, R.J. Lockman, and G.R. Knapp, Ann. Rev. Astron. Astrophys 16, 445, 1978.



Lines in Water Hole      1.42  $\approx$  1.72 GHz      Already Detected Astronomically

<u><math>\nu</math> (GHz)</u>	<u>Molecule</u>	<u>Transition</u>	<u>Reference</u>
1.538 135	NH <sub>2</sub> CHO	$1_{10}^{-1}1_{11}$ F = 1-1	6,7
1.538 693	Formamide	1-2	
1.539 295	Formamide	2-1	
1.539 570	Formamide	1-0	
1.539 851	Formamide	2-2	
1.541 018	Formamide	0-1	
1.570 825	NH <sub>2</sub> <sup>13</sup> CHO	$1_{10}^{-1}1_{11}$ F = 2-2	8
1.610 249	HCOOCH <sub>3</sub> Methyl Formate	$1_{10}^{-1}1_{11}$ A	2,3
1.610 906	Methyl Formate	$1_{10}^{-1}1_{11}$ E	2,3
1.612 2310	OH	$2_{\pi 3/2}$ J=3/2 F=1-2	11,13
1.624 518	<sup>17</sup> OH	$2_{\pi 3/2}$ J=3/2 F, F <sub>1</sub> =7/2, 4-7/2, 4	1,5
1.626 161	<sup>17</sup> OH	$2_{\pi 3/2}$ J=3/2 F, F <sub>1</sub> =9/2, 4-9/2-4	1,5
1.637 564	<sup>18</sup> OH	$2_{\pi 3/2}$ J=3/2 F=1-1	4,6,10,15
1.638 805	HCOOH Formic Acid	$1_{10}^{-1}1_{11}$	17
1.639 503	<sup>18</sup> OH	$2_{\pi 3/2}$ J=3/2 F=2-2	4,6,10,12,15
1.665 4008	OH	$2_{\pi 3/2}$ J=3/2 F=1-1	11,13,14
1.667 3590	OH	$2_{\pi 3/2}$ J=3/2 F=2-2	11,13,14
1.720 5300	OH	$2_{\pi 3/2}$ J=3/2 F=2-1	11,13,14

Table 6.1

Lines in Water Hole of Potential Astrophysical Interest (Not Yet Found)

<u><math>\nu</math> (GHz)</u>	<u>Molecule</u>	<u>Transition</u>	<u>Reference</u>
1.455 7290	$^{17}\text{OH}$	$2\pi_{3/2} \quad J=3/2 \quad F, F_1=5/2, 2-3/2, 1$	1,12
1.474 787	$\text{H}_2\text{CS}$	$10_{2,8}-10_{2,9}$	1,7
1.501 250	$\text{NH}_2\text{CHO}$	$19_{4,15}-19_{4,16}$	1,7
(unsplit - will have hyperfine components)			
1.547 377	$\text{H}_2\text{CO}$	$18_{4,14}-18_{4,15}$	1,7
1.549 798	$\text{H}_2\text{CS}$	$24_{3,21}-24_{3,22}$	1,7
1.584 2740	$^{18}\text{OH}$	$2\pi_{3/2} \quad J=3/2 \quad F=1-2$	1,12
1.596 061	HDS	$3_{3,0}-3_{3,1}$	9
1.604 270	HNCO	$4_{1,3}-4_{1,4}$	10,16
(unsplit - will have hyperfine components)			
1.629 515	$\text{NH}_2\text{CHO}$	$12_{3,9}-12_{3,10}$	1,7
(unsplit - will have hyperfine components)			
1.636 816	HDS	$5_{4,1}-5_{4,2}$	9
1.655 499	$^{17}\text{OH}$	$2\pi_{3/2} \quad J=3/2 \quad F, F_1=5/2, 3-5/2, 3$	12,1
1.656 542	$^{17}\text{OH}$	$F, F_1=7/2, 3-7/2, 3$	12,1
1.659 577	$\text{H}_2^{13}\text{CO}$	$11_{3,8}-11_{3,9}$	1,7
1.683 540	$^{17}\text{OH}$	$2\pi_{3/2} \quad J=3/2 \quad F, F_1=3/2, 1-3/2, 1$	1
1.684 542	$^{17}\text{OH}$	$F, F_1=1/2, 1-1/2, 1$	1
1.692 7950	$^{18}\text{OH}$	$2\pi_{3/2} \quad J=3/2 \quad F=2-1$	1

Only ground vibrational state has been considered here. There are other lines that could be of astrophysical interest in excited vibrational states such as several lines of  $\text{HC}_3\text{N}$ .

Table 6.2

## References

1. Beadett, R.A., and Poynter, R.L., 1978, J. Phys. Chem. Ref. Data, 7, 311.
2. Brown, R.D., Crofts, J.G., Gardner, F.F., Godfrey, P.D., Robinson, B.J., and Whiteoak, J.B., 1975, Ap.J. (Letters), 197, L29.
3. Churchwell, E., and Winnewisser, G. 1975, Astr. and Ap., 45, 229.
4. Gardner, F.F., McGee, R.X., and Sinclair, M.W., 1970, Ap. Letters, 5, 67.
5. Gardner, F.F., and Whiteoak, J.B., 1976, MNRAS, 176, 57P.
6. Gottlieb, C.A., Palmer, P., Rickard, L.J., and Zuckerman, B., 1973, Ap.J., 182, 699.
7. Johnson, D.R., Lovas, F.J., and Kirchhoff, W.H., 1972, J. Phys. Chem. Ref. Data, 1, 1011.
8. Lazareff, B., Lucas, R., and Encrenaz, P., 1978, Astr. and Ap., 70, L77.
9. Lovas, F.J., 1978, J. Phys. Chem. Ref. Data, 7, 1445.
10. Lovas, F.J., and Tiemann, E., 1974, J. Phys. Chem. Ref. Data, 3, 609.
11. Manchester, R.N. and Gordon, M.A., 1971, Ap.J., 169, 507.
12. Rogers, A.E.E., and Barrett, A.H., 1966, Astr.J., 71, 868.
13. ter Meulen, J.J., and Dymanus, A., 1972, Ap.J (Letters), 172, L21.
14. Weinreb, S., Barrett, A.H., Meeks, M.L., and Henry, J.C., 1963, Nature, 200, 829.
15. Wilson, W.J., and Barrett, A.H., 1970, Ap. Letters, 6, 231.
16. Winnewisser, G., Hocking, W.H., and Gerry, M.C.L., 1976, J. Phys. Chem. Ref. Data, 5, 79.
17. Zuckerman, B., Ball, J.A., and Gottlieb, C.A., 1971, Ap. J (Letters), 163, L41.

## **Chapter 7**

### **Geophysical and Material Sciences Applications**

Why do we wish to bear forever the noise  
of these more than another noise so close  
to our dwelling place.

--Robert Frost

## 7.1 - Introduction

The unprecedented combination of broadband coverage and high frequency resolution in the multichannel spectrum analyzer and signal detector provide a means for previously unattainable sensitivity in signal analysis. The criteria used in the choice of nonastrophysical applications for the system are: 1) the signals to be monitored must have a broad total bandwidth, which contains narrowband spectral components of interest, and 2) the signals must have a low signal to noise ratio.

Some of the most promising non-astrophysical applications of the multichannel spectrum analyzer and signal detector are in the field of acoustic or microseismic monitoring. There are currently many ongoing research projects, and numerous diagnostic incipient failure tests and non-destructive tests employing acoustic technology, which are producing a wealth of useful data. The two major impediments in this relatively new and very promising field are frequency resolution and acoustic emissions with low signal to noise ratios. A sensitive signal detector with high frequency resolution might illuminate many of the microseismic or acoustic emission phenomena that are only partially understood at the present time.

Acoustic or microseismic emissions are transient elastic waves generated by rapid release of energy during dynamic processes in materials. Localized transient instabilities such as crack propagation, dislocation motion, and microfracturing in rocks and metals or fiber fracture and crack propagation in polymers and fiber-reinforced plastics (FRP) are examples of the dynamic

processes which produce acoustic emissions. The transient elastic waves or acoustic emissions are detected with piezoelectric transducers which convert them to electrical signals which are then analyzed in a variety of ways which will be discussed later. The sensitivity of acoustic emission analyzers to localized transient instabilities allows prediction of catastrophic failure. Although the stresses in the system are below the elastic limit of the material, the localized stresses in the region around a flaw or crack may be higher than the elastic limit, due to stress concentration, which results in propagation of the flaw or crack. In most material systems catastrophic failures are often the result of numerous localized instabilities which coalesce, and therefore the ability of acoustic emission analysis to detect these localized failures which occur at stresses well below the elastic limit of the material makes prediction of catastrophic failure possible.

In addition to failure prediction, acoustic emission technology has a variety of useful applications. Leaks in pipes, reservoirs, and dams have been successfully detected and located with acoustic emission analysis. Flaws in pressure vessels, reactor vessels, and a variety of other structures have been detected and located with a variety of acoustic non-destructive testing (NDT) methods.

The use of acoustic monitoring was initiated in the late thirties and early forties by Obert (1939, 1940, 1941), Obert and Duvall (1942, 1945, 1945a) Hodgson (1943) and Hodgson and Gibbs (1945), in studies of rock bursts and pillar failures in mines. This early work was crude and employed a low frequency resonant geophone, an audio amplifier, and a headset (audible frequencies were the only frequencies monitored). Kaiser (1950, 1953) and

his coworkers in Germany in the earlier fifties are generally credited with the first modern acoustic studies in work on metals, Kaiser reported that all metals, including steel, zinc, copper, aluminum, and lead produce acoustic emissions under load. He also reported that acoustic phenomena were irreversible, that is, acoustic emissions do not occur in cyclic loading tests unless the load in the cycle exceeds that in all previous loading cycles. This phenomenon is called the Kaiser effect. Schofield (1958) and Tatro (1959) did much to improve the sophistication of acoustic monitoring instrumentation and experimentation in the mid to late fifties with research on the sources of emissions. They found that acoustic emissions in metals were produced by the dislocation movement which accompanies plastic deformation in such materials. Dunne, Tatro, and Harris first investigated the use of instrumentation whose frequency range was well above the audio range. The use of equipment with this range eliminated many of the problems early workers encountered with excessive background ambient noise.

There are numerous present areas of application of acoustic or microseismic analysis. Brittle deformation processes in geologic materials have been greatly illuminated by studies of microseismic emissions. Scholz (1968) proposed that acoustic emissions in polycrystalline geologic materials were produced by microfracturing of individual grains, and that failure in such materials (by coalescence of microfractures) could be predicted on the basis of the acoustic count rate. This work is further supported by Martin and Durham (1975) and Dunning and Dunn (1978), in work on crack propagation in quartz. Lockner, Walsh, and Byerlee (1977) and Gupta (1973) have investigated changes in velocity of acoustic emissions in rocks undergoing deformation. Byerlee and Lockner (1978) and Blacic and Malone (1977) have investigated the

acoustic nature of shear failure and fracture propagation. Byerlee and Lockner (1977) and Hardy (1975, 1978) have found that fluid or gas pressure in geological materials can be estimated on the basis of acoustic activity. The acoustic nature of slope instability in rock, soil, and snow has been investigated by Koerner and Lord (1974, 1975) and Sommerfield (1978). There are many other examples of acoustic applications in geologic problem solving which, for the sake of brevity, cannot be covered in this report.

Acoustic emission techniques are also well-suited to problem solving studies in material research, structural integrity analysis, and other areas of industrial research. Liptai (1972) investigated the acoustic nature of failure in fiberglass-reinforced epoxy. Tetelman, Harris and Darwish (1972) conducted similar experiments with whisker-reinforced composites. Carlyle (1975), Hamstad and Chiao (1976), and Fowler and Gray (1979) have characterized the acoustic nature of failure in fiber-reinforced plastics. Their work indicates that the different mechanisms involved in failure of these materials; such as fiber failure and glass cracking, can be differentiated on the basis of acoustic signature, signal amplitude, and power spectra. Hamstad and Chiao (1976) and Mitchell (1979) have investigated non-destructive tests for fiber-reinforced plastic structures and pressure vessels. Hartbower et al (1972) and Wingfield (1972) have demonstrated that weld-cracking can be detected and located, and that the structural integrity of welds can be tested with acoustic emission techniques. Acoustic emission technology has been applied in the aerospace and aeronautics industries with success by Moore (1970), Green et al (1970), and Nakamura et al (1972). Acoustic analysis of on-line nuclear reactor vessels has been employed by the Connecticut Yankee, and Elk River power plants (Whyllie, 1971). The deformation of concrete and



and other building materials has been characterized acoustically by Li et al. (1970). The mining industry has made extensive and successful use of acoustic monitoring technology. Leighton and Blake (1970) and Krauland and Westerberg (1977) have shown that rock bursts, roof bounces, and pillar failures can be predicted in some instances on the basis of increased acoustic activity prior to the event.

## 7.2 - Acoustic Emission Analysis Techniques

Acoustic activity can be measured and analyzed in a variety of ways. Early workers employed threshold counting techniques, which consist of accumulating a running count of the number of times the amplitude of the signal from the transducer exceeds a predetermined threshold. It has been shown by numerous experimentalists that, in most polycrystalline materials, the count rate (counts per unit time) rises sharply just prior to failure. In some cases the use in count rate can be utilized to predict failure. This technique has some deficiencies, however, which limit its effectiveness. A single acoustic burst of high amplitude may exceed the threshold several times, resulting in a spurious value for total acoustic activity, while acoustic events with lower amplitude produce a single count. Furthermore, in materials that have been loaded in a cyclic manner, the Kaiser effect often precludes prediction of failure if the material fails at a load level attained in a previous cycle.

Another technique commonly employed is burst or event thresholding, which entails accumulating the count rate of acoustic events or bursts above a predetermined threshold. This method has the same limitations as simple thresholding, however the count rate obtained in event thresholding is more accurate and not as sensitive to the level of the threshold.

The maximum signal amplitude during a single event is a commonly employed acoustic analysis technique. Signal amplitude measurements are useful in the laboratory where sample homogeneity is not a problem or in the field when relative changes in amplitude as a function of time for a single transducer are of interest, however in other cases signal amplitude alone is not a particularly informative parameter.

A number of workers have investigated the frequency characteristics and power spectra of individual acoustic events and cumulative acoustic events. Suzuki et al (1964) and Chugh et al (1972) found that acoustic events in rocks undergoing deformation had a regular and predictable power spectra. Chugh et al (1972) found that the frequency of emissions shifts upward as a function of stress level, as the failure stress is approached. Chugh et al propose that spectral data was more useful when presented as an energy distribution ratio which, in that study, was defined as the ratio of acoustic energy in the range 500-5,000 Hz, to that observed in the range of 10,250-15,000 Hz. In general, the energy distribution ratio decreases with increased stress for a variety of rock types. Graham and Alers (1975) found that many metals and composite materials have specific and unique acoustic emission spectra. They also showed that it was possible to detect failure and identify component noises in steam turbine equipment based on the spectra of emissions. Fowler and Gray (1979) and others have shown that fiber failure, crack propagation in glass and other failure mechanisms in fiber-reinforced plastics can be identified by their spectra. A typical frequency-versus-amplitude plot reveals such flaws. The major drawback of spectral analysis is the lack of frequency resolution attainable with available off-the-shelf spectral analysis equipment. Most workers involved with acoustic spectra feel that frequency

resolution is the main impediment to solution of a number of acoustic problems. Many materials have rather wide spectra (up to several MHz) which cannot be observed in total with very high frequency resolution.

Source location analysis is a recently developed acoustic technique that appears to be promising. Source location techniques employ an array of three or more transducers which are located at various points on the body of interest. Acoustic activity from a source is detected by each transducer at a different time depending on the distance from the source to the transducer. A triangulation technique based on the differences in arrival times among the transducers is used to locate the acoustic source empirically. Mogi (1968) and Scholz (1968) have used this technique on a variety of rock types with some success. Byerlee and Lockner (1978) have used the source location technique to map fracture propagation in granite in the most sophisticated geologic source location study yet attempted. Source location techniques are commonly used in industry to identify flaws and fractures (Hamstad and Chiao, 1976; Mitchel, 1979; and Wingfield, 1972). Source location techniques are employed in mines to predict and locate rock bursts, Blake (1977) and Leighton and Steblay (1977). An excellent description of this technique is found in Kelly and Schlamp (1977).

Seismic velocity attenuation analysis is another promising acoustic technique. Lockner et al (1977) have demonstrated that the seismic velocity of acoustic signals decreases up to 30% immediately prior to failure in granite. They attributed this attenuation to the presence of numerous micro-cracks. Hadley (1975), Merkulova et al (1972), and Born (1941) have also investigated this phenomenon with similar results. It has also been observed that seismic velocities are often attenuated in the area of the focus just

the signature of acoustic emissions generated during sliding friction tests on granite was similar to the signature of some earthquakes.

### 7.3 - Possible Applications of the Signal Detector in Acoustic Emission

The two areas of interest in acoustic monitoring which might be greatly illuminated by the multichannel spectrum analyzer and signal detector are spectrum analysis and low SNR signal analysis. Spectrum analysis at the present time is limited by frequency resolution. The acoustic spectra of many materials undergoing deformation have bandwidths of a megahertz or more, while the frequency resolution of presently available off-the-shelf equipment with bandwidths in this range is on the order of 10 kHz. There are undoubtedly numerous narrowband acoustic spectral components that cannot be effectively analyzed with 10 kHz frequency resolution. It can be shown that the effective SNR of a signal whose bandwidth is narrower than the frequency resolution of the analyzer becomes

$$\text{SNR} = \frac{B_s}{F_r \sqrt{F_r}} \quad \text{where } F_r = \text{frequency resolution of the analyzer} \quad (7-1)$$

$B_s = \text{bandwidth of the signal}$

This means that narrowband signals of interest may be lost because of poor frequency resolution. The frequency resolution of the multichannel spectrum analyzer is high enough to effectively investigate narrowband acoustic spectral components.

There are numerous instances, especially in field projects, when the ambient noise is of such great amplitude that many acoustic signals of interest are well below the detection threshold. Acoustic analysis of mines, underground reservoirs, and oil fields is often hindered by high amplitude noise and a low SNR as are laboratory investigations of the Kaiser effect, crack

and fracture propagation, and hydrofracture. The high sensitivity signal detector proposed in this study would greatly enhance our ability to effectively investigate such phenomena. Some specific types of investigations in industry and geophysics that would be greatly enhanced by the multichannel spectrum analyzer and the signal detector will be covered in the next two sections. One limitation of the system, however, is the low time resolution of the spectrum analyzer. Transient emission of much less than a second in duration would not be detected by the signal detector unless the pipeline capability of the spectrum analyzer could be implemented.

#### 7.4 - Industrial Applications

##### Source Mechanism Studies

It has been found by McCabe and Koerner (1979) that the frequency of acoustic emissions produced by crack propagation in coal is shifted downward as the crack propagates, and the source area is enlarged. A rigorous analysis of this apparent phenomenon could result in an effective method for monitoring acoustic sources in order to determine if the flaws producing acoustic signals are growing. The sensitivity and frequency resolution of the proposed signal detection system would be ideal for such an investigation (an experimental procedure outline for a feasibility study of this phenomenon is presented in Appendix A7.1). Source mechanism studies of bearing failure in rotating equipment have so far shown some promise. The spectra of emissions of a bearing undergoing failure can be used to identify the type of bearing, and source location analysis can locate the particular bearing of interest (Pollock, 1979). Source mechanism studies are also of importance in pressure vessels and nuclear reactor vessels. The location, spectral properties and amplitude attenuation of acoustic

signals generated by a flaw or fracture, or the velocity, frequency, and amplitude attenuation (if the coupling parameters are well understood) of induced acoustic signals may allow workers to identify and characterize flaws in such vessels during on-line operation. The signal detection system proposed in this study and the multichannel spectrum analyzer would be capable of such studies.

### Failure Mechanism Studies

The failure mechanisms in metals, plastics, composite materials, and building materials may fall into specific categories based on spectral components, signal amplitude and spectral shift characteristics. It has been shown, for instance, by Fowler and Gray (1979) that fiber failure can be differentiated from glass cracking in fiber-reinforced plastics based on relative amplitudes and spectral components. Further investigation employing a sensitive detector with high frequency resolution may show that intercrystalline fracture, grain boundary deformation, and dislocation movement for various types of materials can be uniquely characterized by the parameters listed previously. High resolution and high sensitivity studies such as these might also reveal that there is a unique frequency component for each material that appears immediately prior to failure, the knowledge of which would greatly enhance the ability to predict failure. It has already been established that the flow mechanisms in plastics can be differentiated and identified by characteristic signatures of acoustic activity in frequency-time space. The signal detection system and multichannel spectrum analyzer proposed in this report could, quite possibly, provide extremely important information on failure mechanism characteristics.

### Leak Detection

It has been demonstrated by Parry (1971) in pressure vessels and pipes and by Koerner and Lord (1974) in earthen dams, that leaks can be detected by analyzing the spectral components and frequency ratios of acoustic emissions. Leaks are often manifested by a single or group of specific spectral components, the characteristics of which are apparently dependent on the pipe, vessel, or dam material and the leaking medium. The signal amplitudes of acoustic emissions resulting from leaks are often quite low, a problem which could be ameliorated to a great extent by the sensitivity and frequency resolution of the proposed signal detection system.

### Kaiser and Felicity Studies

The apparently irreversible nature of acoustic activity was discovered by Kaiser (1950). The Kaiser effect is the immediately irreversible nature of acoustic emission activity resulting from applied stress. If the effect is present, there is little or no acoustic activity until previously applied stress levels are exceeded. This effect is not extremely well-understood for some materials. The Kaiser effect in rocks and concrete may be due to the fact that acoustic emissions are produced by microfracturing, and in cyclic loading new microfractures are not produced until previous loads are exceeded. Previous investigations of the Kaiser effect have employed thresholding techniques which may not have detected low amplitude signals which possibly occur under cyclic loading at stresses below previous stress levels. An investigation of this phenomenon with the multichannel spectrum analyzer and signal detector may uncover low amplitude signals or characteristic spectral components that will enable us to better understand the Kaiser effect.

The felicity effect is a phenomenon which occurs in plastics under cyclic loading. The felicity effect results in significant emissions at stress levels well below previously attained stress levels. It is believed that the felicity effect is caused by isoelastic deformation during unloading which effectively anneals previously formed deformation structures. A highly sensitive high frequency resolution analysis of this phenomenon may reveal the deformation mechanism which produces the felicity effect.

#### Thermoplastic Studies

The failure of thermoplastics and glasses cannot presently be predicted with acoustic monitoring techniques because acoustic emissions are either absent or low enough in amplitude to escape detection. Glass sometimes displays acoustic activity during crack initiation, however, this only occurs when the crack propagation is unstable. The high frequency resolution and sensitivity of the multichannel spectrum analyzer and signal detector proposed in this study may enable detection of acoustic emissions in these materials, if they are present.

There are, undoubtedly, numerous other industrial applications in the field of industrial acoustic emission monitoring for which the signal detection system proposed in this study may be well-suited. The applications discussed in the section are, however, of particular interest in this field.

#### 7.5 - Geophysical Applications

##### Fracture Density Studies

Fracture density is of great interest to geophysicists, structural geologists and petroleum geologists. Fracture density in fault zones appears



to be an important parameter in fault mechanics, a clear understanding of which may lead to significant advances in the field of earthquake prediction. The density and nature of fractures in petroleum reservoirs are two of the most important factors in effective tertiary recovery of oil in fractured reservoirs.

Present fracture density studies are implemented in a variety of ways. Many geophysicists and petroleum geologists investigate seismic velocity attenuation. In this test an ultrasonic or shot blast signal is activated and the travel times of the elastic waves from the shot site to the sites are recorded and compared to the predicted travel times for the unfractured rock type of interest. If the fracture density is high, the velocity of the elastic waves will be greatly attenuated from the predicted value. This method has some limitations, most notably sensitivity to fracture filling. A fracture which is open and filled with gouge or fluid will produce the same velocity attenuation as numerous closed fractures. Another parameter measured by geophysicists and petroleum geologists and engineers is the Q factor which is basically a measure of the energy loss of an input ultrasonic or shot signal. A highly fractured body of rock will have a lower Q value than unfractured rock of the same composition. This method has the same basic limitations as velocity attenuation because the parameters used to calculate both are basically the same. Measurement of amplitude attenuation as a function of fracture density has been attempted in the laboratory, however the uncertainty involved in determining the coupling parameters of the transducer and the rock makes field utilization of this technique somewhat difficult. Power spectral analysis has been employed with some success in recent years. It has been found that high fracture density results in a downward frequency shift of the power spectrum.

The frequency and amplitude resolution of the multichannel spectrum analyzer and signal detector allow more detailed and sophisticated analyses of fracture density. Wider bandwidth input signals can be employed and frequency attenuation, amplitude attenuation (if the coupling parameters are known), and detailed power spectra measurements would be possible. An analysis of  $Q$  as a function of frequency might also be productive because Aki (1979) has found that  $Q$  varies as a function of frequency to some extent, in addition to varying as a function of fracture density.

### Stress Field Analysis

Accurate measurements of stress fields and stress orientations are of great importance to geophysicists, petroleum engineers, and mining engineers. A sudden change in the stress field of a fault zone may be a premonitory event, or an unexpected variation of the stress field with depth may result in expensive drilling equipment damage in an oil well. Rock bursts and other mine failures are often preceded by rapid variation in stress field intensity and orientation. Presently such techniques as overcoring, hydrofracture strength analysis, and extensometer and strain gage analysis are used to measure stress field orientation and magnitude. An interesting application of the multichannel spectrum analyzer and signal detector would be an analysis of the acoustic emission activity, under loading, of cores cut from fault zones, mine walls, or drill cores in order to determine the stress field by investigation of the Kaiser effect. The Kaiser effect could be used as a stress level indicator of great sensitivity. One major limitation of the use of the Kaiser effect to determine stress levels is the fact that the stress values calculated by determining the stress level at which acoustic activity

begins may be higher than the actual stress level, if the stress level in the rock was significantly higher at some time shortly before the cores were cut. Kaiser effect studies should, however, be undertaken because the spectrum analyzer and signal detector are well-suited for this type of investigation and Kaiser effect measurements may prove to be an excellent and inexpensive measurement technique. An experimental procedure outline for a feasibility study of an investigation of Kaiser effect as a stress indicator is presented in Appendix A7.1.

#### Source Mechanism Studies

It has been established that the source mechanism of acoustic emissions in metals and plastics can be determined on the basis of frequency and amplitude. Fiber fracture, dislocation movement, crack propagation, and grain boundary deformation all have different amplitudes and spectral components. It is probable that the same situation exists in rocks; source mechanisms such as dislocation movement, microfracturing, grain boundary deformation, and inter-crystalline crack propagation should be distinguishable on the basis of spectral components and relative signal amplitudes. This has not been investigated as of yet because the spectrum of acoustic emissions in rocks is wide enough in bandwidth to preclude high resolution spectral analysis with available equipment.

Another type of source mechanism study that would be productive is spectral analysis of a propagating flaw or fracture. McCable and Koerner (1979) have found that the spectrum of emissions from a propagating cleat in coal shifts downward in frequency as the cleat dimensions increase. This apparent phenomenon should be investigated with a high resolution detection

system such as the one proposed in this study because if this phenomenon exists and can be clearly understood, significant progress may be made in prediction of pillar failures and rock bursts in mines.

#### General Deformation Studies

The acoustic nature of deformation processes in rocks is not very well understood at the present time due in great part to the broad bandwidth of the emissions and the lack of frequency resolution of existing acoustic monitoring equipment. Numerous types of investigations of rock deformation could be carried out with the signal detection system proposed in this report. Compressive and tensile failure tests might illuminate the acoustic nature of the shift from microfracturing to cataclastic failure in rocks. The acoustic nature of hydrofracturing could well be delineated in a more rigorous manner with such a signal detector. Cyclic loading tests might illuminate the acoustic nature of failure under cyclic loading. There are undoubtedly numerous additional studies that could be implemented with the multichannel spectrum analyzer and signal detector proposed in this study that would greatly enhance our knowledge of deformation processes in rocks.

### References

1. Aki, K., 1979, Personal Communication, Mass. Inst. of Tech., Cambridge, Mass.
2. Blacic, J., and Malone, S., 1977, "Preliminary Analysis of Micro-Seismic Signals Associated with Rock Fracture," Geophys. Res. Letters, Vol. 4, pp. 477-479.
3. Blake, W., 1977, "Design Installation and Operation of Computer Controlled Rockburst Monitoring Systems" in Proc. of the First Conference on Acoustic/Microseismic Activity in Geologic Structures and Materials. Trans. Tech. Press., Aedermannsdorf, Switz.
4. Born, W., 1941, The Attenuation Constant of Earth Materials, Geophysics, Vol. 6, pp. 132-148.
5. Byerlee, J. and Lockner, D., 1977, "Acoustic Emission during Fluid Injection into Rock" in Proceedings of the First Conference on Acoustic Emission in Geologic Structures and Materials, Penn. State University, Trans. Tech. Press, Aedermannsdorf, Switz.
6. Carlyle, J., 1975, "Acoustic Emission in Fiber-Reinforced Composites," Dept. of the Navy Report # NADC-75082-30.
7. Chugh, Y., Hardy, H., and Stefanko, R., 1972, "An Investigation of the Frequency Spectra of Microseismic Activity in Rock under Tension", Proc. of the Tenth Rock Mechanics Symposium, pp. 73-113, AIME, New York.
8. Dunegan, H., Tatro, C., and Harris, D., 1964, "Acoustic Emission Report" Report UCID-4868, Revision 1, Lawrence Radiation Lab., Livermore, CA.
9. Dunning, J., and Dunn, D., 1978, "Microseismicity of Stable Crack Propagation in Quartz" in Proceedings of the Second Conference on Acoustic/Microseismic Activity in Geologic Structures and Materials, No. 18-23 1978, Penn. State University, Trans. Tech. Press, Aedermannsdorf, Switz.
10. Fowler, T., and Gray, E., 1979, "Development of an Acoustic Emission Test for FRP Equipment", ASCE Preprint 3583.
11. Graham, L., and Alers., G., 1975, "Acoustic Emission in the Frequency Domain", ASTM Special TEchnical Publication 571, 1975, Spanner and Mcelroy (Eds.), pp. 11-39.
12. Green, A., Dunegan, H., and Tetelman, A., 1973, "Nondestructive Inspection of Aircraft Structures and Materials via Acoustic Emission", Technical Report DRC-107, Dunegan Research Corp., Livermore, CA.
13. Gupta, I., 1973, "Seismic Velocities in Rock Subjected to Axial Loading up to Shear Fracture", Journal of Geophysical Research, Vol. 78, pp. 6938-6942.

14. Hadley, K. 1975, "Further Studies in Crystalline Rock", Ph.D. Thesis, Mass. Inst. of Technol., Cambridge, Mass.
15. Hamstad, M., and Chiao, T., 1976, "Structural Integrity of Fiber/Epoxy Vessels by Acoustic Emission", SAMPE Quarterly Report, October, 1974.
16. Hardy, H.R., 1975, "Application of Acoustic Emission in the Evaluation of Underground Gas Storage Reservoir Stability", Proceedings of the Ninth Canadian Symposium on Rock Mechanics, Montreal 1973, pp. 77-111.
17. Hardy, H.R., 1978, "Some Current Applications of Microseismic Techniques in Geomechanics" in Proceedings of the Advanced Study Institute on Dynamical Methods in Soil and Rock Mechanics, Karlsruhe W. Germany, A.A. Balkema Press.
18. Hartbower, C., Reuter, W., Morais, C., and Crimmins, P., 1972, "Use of Acoustic Emission for the Detection of Weld and Stress Corrosion Cracking", Acoustic Emission ASTM Specail Technical Publication 505, pp. 187-221.
19. Hodgson, A., 1943, Canadian Institute of Mining and Metallurgy Transactions, TCIMA, Vol. 46, 1943, pp. 313-324.
20. Hodgson, A., and Gibbs, Z., 1945, "Seismic Research Program--Rock Burst Problem, Lakeshore Mines", Dept. of Mines, Resources, and Surveys, Engineering Branch Report 14, Ottawa, Canada.
21. Kaiser, J., 1950, "Untersuchungen Über das Auftreten Gerauschen Bein Zugversuch", Ph.D. thesis, Technische Hochschule, Munchen.
22. Kaiser, J., 1953, "Untersuchungen Über das Auftreten Gerauschen Beim Zugversuch", Arkiv Fur Das Eisenhüttenwesen, AREIA, Vol. 24, No.1/2 Jan/Feb, pp. 43-45.
23. Kanamori, H., and Fuis, G., 1975, "Variation of P Wave Velocity before and after the Galway Lake Earthquake" (abstract), EOS Trans. AGU, 56, p. 1019.
24. Kelly, M., and Schlamp, R., 1977, "Detecting Structural Degradation by Acoustic Emission", Dunegan-Endevco Technical Report, DE 76-11, Dunegan Endevo Corp., San Juan Capistrano, CA.
25. Koerner, R., and Lord, A., 1974, "Earth Dam Warning System to Prevent Hazardous Material Spills", AIChE/EPA Conference on Control of Hazardous Material Spills, San Francisco, CA.
26. Koerner, R., and Lord, A., 1974, "Acoustic Emission in Stress Soil Samples", J. Acoust. Soc. of Am., Vol. 56, No. 6, pp. 1924-1927.

27. Koerner, R., and Lord, A., 1975, "Acoustic Emission Monitoring of the Stability of Earthen Dams", Int'l. Water Power and Dam Construction, Vol. 27, No. 9.
28. Krauland, N., and Westerberg, K., 1977, "Experiences with AE-Measurement in Boliden Mines", Proceedings First Conference on Acoustic Microseismic Activity in Geologic Structures and Materials, Penn. State Univ., Trans. Tech., Press, Aedermannsdorf, Switz.
29. Leighton, F., and Blake, W., 1970, "Rock Noise Source Location Techniques", U.S. Bureau of Mines R.I. 7432.
30. Leighton, F., and Steblay, B., 1977, "Applications of Microseismics in Coal Mines" in Proc. of the First Conference on Acoustic/Microseismic Activity in Geologic Structure and Materials, Trans. Tech. Press, Aedermannsdorf, Switz.
31. Li, S., Ramakrishnan, V., and Russel, J., 1970, "Where Stands Nondestructive Testing of Concrete and Whither?" Int. Journ. of Nondestructive Testing, Vol. 2, pp. 281-300.
32. Liptai, R., 1972, in Composite Materials: Testing and Design. ASTM Special Technical Publication, 497, p. 285.
33. Lockner, D., Walsh, J., and Byerlee, T., 1977, "Changes in Seismic Velocity and Attenuation during Deformation of Granite", Journal of Geophysical Research, Vol. 82, pp. 5374-5378.
34. Lockner, D., and Byerlee, J., 1978, in Development of Fracture Planes During Creep in Granite. Second conference on Acoustic/Microseismic Activity in Geologic Structures and Materials, Penn. State University, Trans. Tech. Press, Aedermannsdorf, Switz.
35. Mantis, C., and Lindh, A., 1976, "The Oroville Foreshocks and an Apparent Coseismic Change in Fault Plane Orientation with Short-Term Precursor" (abstract), EOS Transactions of AGU, 57, pp. 956.
36. Martin, R., and Durham, W., 1975, "Mechanisms of Crack Growth in Quartz" Jour. of Geophys. Res., Vol. 80, pp. 4837-4844.
37. McCabe, M., and Koerner, R., 1979, Unpublished data, personal communication, Drexel University, Philadelphia, PA.
38. Merkulova, V., Pigulerskiy, and Isaplev, V., 1972, "Sound Absorption Measurements in Uniaxially Compressed Rocks", Izv. Acad. Sci U.S.S.R. Phys. Solid Earth, Vol. 3, pp.166-167.
39. Mitchell, J., 1979, "Fundamentals of Acoustic Emission Application as a NDT Tool for FRP", Presented at the 34th Annual Technical Conference, Reinforced Plastics/Composites Institute, The Society of the Plastics Industry, Inc., New Orleans, Jan. 1979.
40. Mogi, K., 1968, "Source Location of Elastic Shocks in the Fracturing Process in Rocks", Bull. Earthquake Res. Institute, Vol. 46, pp.1103-1125.

41. Moore, J., 1970, "Early Detection of Fatigue Damage", NA-70-640, 3rd Semiannual Report, North American Rockwell Corp, Los Angeles.
42. Nakamuray, McCavley, B., and Veack, C., 1972, "Study of Acoustic Emission during Mechanical Tests of Large Flight Weight Tank Structure", MSC-04800 FZ K-390, General Dynamics, Convair Aerospace Divison. also in Proceedings of the 9th symposium on Nondestructive Emission testing, San Antonio, 1973.
43. Nur, A., 1972, "Dilatency, Pore Fluids, and Premonitory Variations of  $t_s/t_p$  Travel Times", Bull Seism. Soc. Amer., Vol. 62, pp. 1217-1222.
44. Obert, L., 1939, "Measurement of Pressure on Rock Pillars in Underground Mines", U.S. Bureau of Mines RI 3444 Part I.
45. Obert, L., 1940, "Measurement of Pressure on Rock Pillars in Underground Mines", U.S. Bureau of Mines RI 3444 Part II.
46. Obert, L., 1941, "Use of Subaudible Noise for Prediction of Rock Bursts", U.S. Bureau of Mines RI 3555.
47. Obert, L., and Duvall, W., 1942, "Use of Subaudible Noise for Prediction of Rock Bursts Part II", U.S. Bureau of Mines RI 3654.
48. Obert, L., and Duvall, W., 1945, "Microseismic Method of Predicting Rock Failure in Underground Mining, Part I: General Methods", U.S. Bureau of Mines RI 3797.
49. Obert, L., and Duvall, W., 1945, "Microseismic Method of Predicting Rock Failure in Underground Mining, Part II: Laboratory Experiments", U.S. Bureau of Mines RI 3803.
50. Parry, D., 1971, "Acoustic Emission Analysis of Pressure Vessels and Piping", Proceedings of the 1st International Conference on Structural Mechanics in Reactor Technology, E UR 4820, Vol. 4, Commission of the European Communities, pp. 553-568.
51. Pollock, A., 1979, Personal communication, Dunegan/Endevco Corp., San Juan Capistrano, CA.
52. Tatro, C., 1959, "Sonic Techniques in the Detection of Crystal Slip in Metals", Status report, Division of Engineering Research, College of Engineering, Michigan State University, East Lansing, Michigan.
53. Tetelman, A., Harris, D., and Harwish, F., 1972, in Acoustic Emission ASTM Special Technical Publication 505, pp. 238-249.
54. Schofield, B. et al. 1958. "Acoustic Emission under Applied Stress" ASTIA Document AD 155 674, WADC Technical Report, pp. 58-194.



55. Scholz, C., 1968, "Experimental Study of the Fracturing Process in Brittle Rock", Journal of Geophysical Research, 73, pp. 1447-1452.
56. Sommerfield, R., 1978, "Acoustic Emissions from Unstable Snow Slopes", Proceedings of the Second Conference on Acoustic/Microseismic Activity in Geologic Structures and Materials, Penn. State University, Trans. Tech. Press.
57. Suzuki, T., et al., 1964, "A New Approach to the Prediction of Failure by Rock Noise", Fourth Inter. Conference on Strata Control and Rock Mechanics.
58. Wylei, R.D., 1971, "In-Service Inspection Program for Nuclear Reactor Vessels", SWRI Project 17-2440, Biannual Progress Report #4, Southwest Research Institute, San Antonio, TX, June 1971.
59. Wingfield, P., 1972, "Weld Quality Monitoring by Stress Wave Emission", Welding and Metal Fabrication, WLMFA, Vol. 40, pp. 351-355.

#### General References

60. Proceedings of the First Conference on Acoustic/Microseismic Activity in Geologic Structures and Materials, Trans. Tech. Press, Aedermannsdorf, Switz., 1979.
61. Proceedings of the Second Conference on Acoustic/Microseismic Activity in Geologic Structures and Materials, Trans. Tech. Press, Aedermannsdorf, Switz., 1979.
62. Acoustic Emission, ASTM Special Technical Publication 505, 1972, Liptai, Harris, and Tatro (eds.).
63. Monitoring Structural Integrity by Acoustic Emission, ASTM Special Technical Publication, 571, 1975, Spanner and McElroy (eds.).

Postlude

The operator powers up the Oasis system, mounts fresh tapes on the replay/archive drives, and places the system in self-diagnostic mode. Momentarily, the status display reports that all subsystems are functioning properly.

Via a magnetic tape, the operator enters into the RFI catalog data concerning astrophysical signals in the portion of sky to be covered today. He now activates the local RFI receiver which begins to dynamically update the catalog.

Pointing the telescope at a known quiet portion of the sky, the operator places Oasis into threshold-setting mode. The control CPU makes minor adjustments to thresholds in the carrier wave detector, pulse detector and cluster detector systems.

On this particular day, a radio astronomer is on hand to continue a high resolution OH survey along the lines of sight to be searched. She takes her seat at the radio astronomy terminal and loads a program on her minicomputer.

The SETI operator, meanwhile, is specifying to the control CPU the protocol for making automatic entries to the system log and event archives, for generating graphic displays and audible alarms.

When all is ready, he points the antenna to the first star (and any unseen surrounding system of planets) and the search begins.

Digital data from multi-channel spectrum analyzers (MCSAs) connected to two orthogonally-polarized receiving systems commences to flow at the rate of 8 million complex spectral values per receiver per second in four parallel data streams into the NBIT detector, the pulse detector, the

carrier wave detector, and a 100 giga-bit tape memory. In the NBIT system the stream divides again to supply 12 processors each at the same data rate: 32 megawords per second. The NBIT low-threshold feature detectors are now providing the clustering system's histogram unit with numerous patterns to count. The actual cluster search will begin at the end of the 1000-second observation of this star.

At the first 20-second decision point, the carrier wave processor makes several detections, but the control CPU matches these with known military radar signals in the RFI catalog so no further action is taken on them. Nor does anything unusual appear in the coarse-resolution spectrum the operator has chosen for his display.

Two minutes into the search, the audible alarm sounds and an integrated coarse spectrum of somewhat different resolution appears on the graphics display. A small but noticeable peak is seen in the middle of the spectrum. The status display indicates that the binning system has made a detection and the moving system log and event archive tapes indicate that information is being written on them.

It looks like a maser--a previously unknown one. "Do you have a maser on your screen?" he asks the radio astronomer. "Yes I do, now," she replies. "You should record it." "It already has been."

He enters a comment in the system log. Over the remainder of the observation of the star the maser line shows more and more clearly. The operator elects to save the entire record on the total observation archive.

The next few stars reveal nothing unusual. Later in the afternoon, the clustering algorithm finds some extremely unusual concentrations of broadband pulse features in an observation of Epsilon Eridani. The SETI

operator saves the total observation for further analysis, along with an image of the cluster as it appears on the display.

Detection is, or ought to be, an exact science, and should be treated in the same cold and unemotional manner. You have attempted to tinge it with romanticism, which produces much the same effect as if you worked a love-story or an elopement into the fifth proposition of Euclid.

--A.C. Doyle, The Sign of Four

## Appendix A2.1

### The Discrete Fourier Transform of Signals and Noise

#### The Discrete Fourier Transform

Assume that we have some time process  $x(t)$ , either deterministic or stochastic, which is sampled every  $T/N$  seconds where  $T$  is the length of one complete record or realization, and  $N$  is the number of samples per record. This is illustrated in Fig. A2.1.1. (In the MCSA which is a part of the system under study here, two quadrature time signals are sampled in this way.)

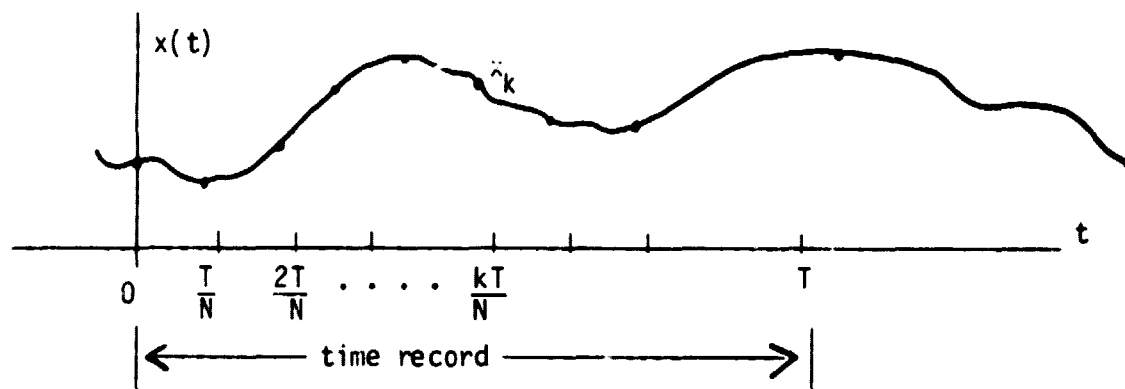


Figure A2.1.1 - Sampling a Continuous Time Function

The Discrete Fourier Transform is defined as:

$$X_n = \frac{1}{N} \sum_{k=0}^{N-1} x_k e^{-j2\pi nk/N}, \quad n = 0, 1, 2, \dots, N-1 \quad (\text{A2.1-1})$$

If the  $x_k$  are samples of a real baseband process with a maximum

frequency  $w$ , the sampling theorem requires that the sampling frequency be at least  $2W$ , or that in  $T$  seconds the number of samples be at least  $2WT$ . If the  $x_k$  are samples of a baseband process obtained from a bandpass process of bandwidth  $B_T$ , the number of samples required in time  $T$  is at least  $2B_T T$  but these samples are of the real and imaginary parts of the complex envelope of the bandpass process, where these parts are sampled once every  $B_T^{-1}$  seconds. This is called quadrature sampling (Peebles, 1979).

The bandpass signal is detected by multiplying it in parallel channels by  $\cos W_c t$  and  $\sin W_c t$ , where  $W_c$  is a constant angular frequency, and thus yielding the real and imaginary parts of the complex envelope. These two signals are low pass filtered and then sampled to produce the real and imaginary parts of the complex samples  $x_k$  used in Equation (A2.1-1).

The values of  $X_n$  generated by Equation (A2.1-1) are complex numbers corresponding to the signal which is in each of  $N$  frequency bins. (If, for example, a bandpass signal of bandwidth 10 MHz is to be divided into one million 10 Hz channels, one million samples will be taken of the real and imaginary parts of the complex envelope, and one million values of  $X_n$  will be obtained ( $N = 10^6$ ).)

The bandwidth of each individual bin is:

$$B = B_T/N \quad (A2.1-2)$$

The center frequency of the  $n^{\text{th}}$  bin is:

$$f_n = \frac{nB_T}{N} = nB = \frac{n}{NT_S}, \quad n = 0, 1, 2, \dots, N-1 \quad (A2.1-3)$$

where  $T_S = T/N$  is the sampling time which equals  $B_T^{-1}$ . Hence the output of the DFT is a set of  $N$  complex numbers each one of which is associated with a frequency bin with bandwidth  $B$  and center frequency  $f_n$ . The baseband representation of this result is shown in Figure A2.1.2.



Figure A2.1.2 - Discrete Fourier Transform Frequency Bins

### Response of the DFT to Sinusoid Signals

Next we consider the case where the input  $x(t)$  to the DFT is a complex sinusoid of the form

$$x(t) = Ae^{jW_r t} \quad (A2.1-4)$$

where  $W_r$  is a fixed baseband frequency. A signal of this form corresponds to a sinusoidal signal in the original passband. In order to conveniently relate cyclic frequency  $f_r$  ( $= W_r/2\pi$ ) to the frequency bins shown in Fig. A2.1.2, we define  $f_r$  in terms of the binwidths  $B$  as:

$$f_r \triangleq rB \quad (A2.1-5)$$

Hence, for example,  $f_r$  lies within Bin 2 whenever  $1.5 < r < 2.5$ . We have then:

$$x(t) = Ae^{j2\pi r B t} \quad (A2.1-6)$$

Since the  $k^{\text{th}}$  sample is taken at the time  $t = kT_s$ , the  $k^{\text{th}}$  sample of  $x(t)$  is:

$$x_k = Ae^{j2\pi r B k T_s} \quad (A2.1-7)$$

Using Equations (A2.1-3) and (A2.1-5),  $x_k$  can be written as:

$$x_k = Ae^{j\frac{2\pi}{N} k r} \quad (A2.1-8)$$

If this sampled signal is now inserted into the DFT, Equation (A2.1-1) becomes:

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} Ae^{j\frac{2\pi}{N} k(r-n)} \quad (A2.1-9)$$



Consider first the simple example where  $r = n$ , corresponding to an  $f_r$  in the center of bin  $n$ . The effect of the summation in Equation (A2.1-9) is simply to add  $A$  a total of  $N$  times so that  $X_n$  equals  $A$ . Thus the transform has a "gain" of unity in center band.

Next consider a signal  $x(t)$  which has the center frequency of Bin  $n$  but which has a phase angle  $\theta$ . That is:

$$x(t) = Ae^{j(W_n t + \theta)} \quad (\text{A2.1-10})$$

Then, clearly:

$$X_n = Ae^{j\theta} \quad (\text{A2.1-11})$$

Hence  $X_n$ , the complex number out of the DFT has the same amplitude for an input signal of any phase, and has the same angle as that of the input.

Let us now generalize to the case where  $r$  takes some general value, not necessarily equal to  $n$ . Define  $d$  to be the difference between  $r$  and  $n$ :

$$d \triangleq r - n \quad (\text{A2.1-12})$$

Assume now that the input signal is:

$$x(t) = Ae^{j\theta} e^{jW_r t} \quad (\text{A2.1-13})$$

$$x_k = Ae^{j\theta} e^{j\frac{2\pi}{N} kr} \quad (\text{A2.1-14})$$

so that the DFT is:

$$X_n = \frac{1}{N} Ae^{j\theta} \sum_{k=0}^{N-1} e^{j\frac{2\pi}{N} kd} \quad (\text{A2.1-15})$$

$$= \frac{Ae^{j\theta}}{N} \sum_{k=0}^{N-1} e^{j\frac{2\pi d}{N} k} \quad (\text{A2.1-16})$$

and since  $\sum_{k=0}^K a^k = \frac{1 - a^{K+1}}{1 - a}$ , (A2.1-17)

$$X_n = \frac{Ae^{j\theta}}{N} \frac{1 - e^{j2\pi d}}{1 - e^{j2\pi d/N}} \quad (\text{A2.1-18})$$

$$= \frac{Ae^{j\theta}}{N} e^{j\pi(d-d/N)} \frac{\sin \pi d}{\sin \pi d/N} \quad (\text{A2.1-19})$$

If  $N$  is large, which is our case,  $X_n$  can be approximated by:

$$X_n \approx Ae^{j\theta} e^{j\pi d} \text{sinc } d \quad (\text{A2.1-20})$$

where:

$$\text{sinc } x \triangleq \frac{\sin \pi x}{\pi x} \quad (\text{A2.1-21})$$

Hence the magnitude and phase of  $X_n$  are respectively:

$$|X_n| = A \text{ sinc } d \quad (\text{A2.1-22})$$

$$\angle X_n = \theta + \pi d \quad (\text{A2.1-23})$$

These outputs (response characteristics) are plotted in Figure A2.1.3 for  $\theta$  equal to 0.

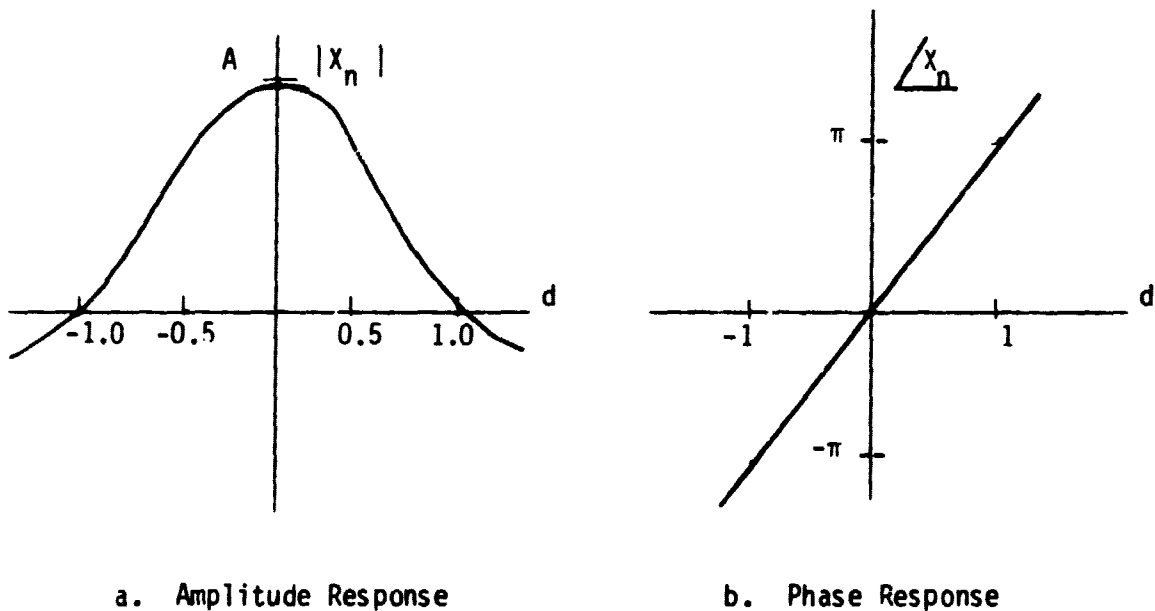


Figure A2.1.3- Response of DFT to Sinusoidal Signal

It is apparent from Equations (A2.1-22) and (A2.1-23) and Figure A2.1.3 that the response of the DFT to a sinusoid which is not in the center of the bin will be less than unity in amplitude, and will introduce some phase.

We consider next the relationship between  $X_n$  as obtained from one set of samples (one record) of  $x(t)$  and the  $X_n$  obtained from the next record, which is obtained  $T$  seconds after the first. In Equation (A2.1-7) substitute  $kT_s + T$  for  $kT_s$ .

$$\begin{aligned}
 x_k &= Ae^{j2\pi rB(kT_s + T)} \\
 &= Ae^{j2\pi rBkT_s} e^{j2\pi rB T} \\
 &= Ae^{j2\pi rBkT_s} e^{j2\pi(d + n)} \\
 &= Ae^{j2\pi rBkT_s} e^{j2\pi d}
 \end{aligned} \tag{A2.1-24}$$

since  $BT = 1$  and  $e^{j2\pi n} = 1$ .

Since the values of  $x_k$  in Equations (A2.1-7) and (A2.1-24) differ only by the term  $e^{j2\pi d}$  which does not depend on  $k$ , the resulting DFT's will differ only by this factor. Thus the angle of  $X_n$  obtained from the second record will differ from that of the first by  $2\pi d$ .

The implication of the above result is that any phase coherence which the sampled signal possesses will be lost in the DFT operation if the signal is not in the center of the band. If, however, an  $X_n$  is observed which changes phase angle linearly as a function of time but does not change in amplitude, then it can be deduced that the input signal is a sinusoid of frequency

$$f_r = f_n + \frac{TB\Delta\theta}{2\pi} \tag{A2.1-25}$$

where  $\Delta\theta$  is the change in the phase of  $X_n$  over  $T$  seconds, that is from one record to the next. A first order indication of a linear change in phase is

an equal phase difference between successive pairs of  $X_n$ 's. That is:

$$\theta_2 - \theta_1 = \theta_1 - \theta_0$$

$$\theta_2 - 2\theta_1 + \theta_0 = 0 \quad (\text{A2.1-26})$$

The test implied by Equation (A2.1-26) could be implemented quite easily to search for constant frequency signals as processed by a DFT.

We consider next the case of a sinusoid which drifts linearly in frequency as a function of time. That is, the instantaneous frequency is:

$$W_i = W_r + \dot{W}t \quad (\text{A2.1-27})$$

where  $\dot{W}$  is the drift rate in radians per second. The corresponding signal is:

$$x(t) = Ae^{j(\dot{W}t^2/2 + W_r t + \theta)} \quad (\text{A2.1-28})$$

Three cases will be examined depending on the value of  $\dot{W}$ .

#### Case 1: Small $\dot{W}$

We assume first that the instantaneous frequency is nearly constant within the bin of interest over one record, of length  $T$  seconds. We express this condition as:

$$\frac{W}{2\pi}T \ll B \rightarrow W \ll \frac{2\pi B}{T} \quad (\text{A2.1-29})$$

In this case the frequency input is essentially constant over one set of data. If that frequency is  $W_r$ , the output  $X_n$  is approximately equal to that given by Equation (A2.1-19).

## Case 2: Intermediate $\dot{W}$

We start by assuming that the record begins at  $t = 0$  so that the  $k^{\text{th}}$  sample is again taken at  $t = kT_s$ . The sampled input signal is then:

$$x_k = Ae^{j\theta} e^{j(W_r k T_s + W k^2 T_s^2 / 2)} \quad (\text{A2.1-30})$$

and hence:

$$X_n = \frac{Ae^{j\theta}}{N} \sum_{k=0}^{N-1} e^{j\left(\frac{2\pi d}{N} k + \frac{\dot{W}}{2B^2 N^2} k^2\right)} \quad (\text{A2.1-31})$$

The effect of the second term (due to frequency drift) in the exponent is to shift each of the  $k$  complex terms in the sum by the angle  $\dot{W} k^2 / 2B^2 N^2$ . If this shift is small compared with  $2\pi d k / N$ , we can neglect it, and we are then back to Case 1. But now we are considering the case where this term is too large to neglect. The summation in Equation (A2.1-31) is not tractable. Our approach is to obtain an approximation by calculating  $X_n$  for a number of fixed values of frequency over which the signal drifts in  $T$  seconds, and averaging over these  $X_n$ . That is, since  $X_n$  for a fixed frequency is given by Equation (A2.1-20),  $X_n$  for a drifting frequency is approximated by

$$X_n \approx \frac{Ae^{j\theta}}{M} \sum_{m=0}^{M-1} e^{j\pi d_m} \text{sinc } d_m \quad (\text{A2.1-32})$$

Where:

$$d_m = d_0 + \frac{\dot{W}}{2\pi B} \frac{mT}{M-1} \quad (\text{A2.1-33})$$

As an example to illustrate this approach consider a signal whose frequency drifts from  $f_n$  ( $d = 0$ ) at  $t = 0$  to  $f_n + B$  ( $d = 1$ ) at  $t = T$ . Let  $M = 3$  so that the fixed-frequency  $X_n$  are obtained for  $d = 0, 1/2$ , and  $1$ . These values are:

$$X_{n_0} = Ae^{j\theta}$$

$$x_{n_1} = Ae^{j\theta}(+j0.64)$$

$$x_{n_2} = Ae^{j\theta}(0)$$

and the average is:

$$\bar{x}_n = 0.33 + j0.21$$

This procedure was repeated for a number of values of M with the results shown in Table A2.1.1.

<u>M</u>	<u><math>\bar{x}_n/Ae^{j\theta}</math></u>
4	0.3017 + j0.2686
8	0.2615 + j0.3348
12	0.2492 + j0.3536
24	0.2373 + j0.3713
100	0.2284 + j0.3840

Table A2.1.1 - Approximate  $\bar{x}_n$  for  $0 \leq d \leq 1.0$

The results approach an asymptotic value. We shall accept the value for M = 100. That is:

$$x \approx Ae^{j\theta}(0.23 + j0.38)$$

$$= Ae^{j\theta}0.44/58.8^\circ$$

This suggests that a drifting signal is definitely attenuated by a significant amount if it drifts from center bin to an edge in one observation time.

We next consider the case for a signal which drifts from the center of a bin to a point part way to the edge ( $0 \leq d \leq a$ ). The resulting approximate

values of  $\bar{X}_n$  were calculated by an averaging process paralleling that suggested by Eq. (A2.1-32). The resulting values of  $X_n$  are given in Table A2.1.2 for four values of  $a$ .

<u>Drift Range</u>	<u><math>X_n/Ae^{j\theta}</math></u>
$0 \leq d \leq 1.0$	$0.44/58.8^\circ$
$0 \leq d \leq 0.5$	$0.79/41.4^\circ$
$0 \leq d \leq 0.2$	$0.96/16.9^\circ$
$0 \leq d \leq 0.1$	$0.99/8.1^\circ$

Table A2.1.2 - Approximate  $X_n$  for Various Drift Ranges

These results show that the DFT can accommodate a certain amount of drift per record, up to perhaps 20% of the binwidth, without serious degradation. However, a drift from the center of the bin to an edge clearly produces a very significant attenuation in the DFT output. Clearly it is desirable that the relative drift per record (compared with the bandwidth) be kept small to maintain the basic system sensitivity. Two strategies are available to keep the relative drift small. These are:

- a. Increase the binwidth  $B$
- b. Change the center frequency of the bin,  $f_n$ , as the signal drifts.

The first strategy has the disadvantage that it increases the amount of noise passed by the bin. The second strategy has the disadvantage that it requires a method of forcing  $f_n$  to track the drifting input signal.

### Case 3: Large $\dot{W}$

Next we consider the case in which  $\dot{W}$  is very large. Specifically:

$$\frac{\dot{W}}{2\pi}T \gg B \implies \dot{W} \gg \frac{2\pi B}{T} \quad (\text{A2.1-34})$$

This means that the carrier signal sweeps through the bin in a time which is short compared with the record time  $T$ .

The problem is simplified by assuming that the response is due only to the input while the carrier is within the bin (i.e.  $-1 \leq d \leq 1$ ). This is equivalent to assuming in Figure A2.1.3a that  $X_n = 0$  when  $|d| > 1$ . In this case  $X_n$  can be approximated by:

$$\bar{X}_n \approx \frac{Ae^{j\theta}}{2M-1} \sum_{m=-M+1}^{M-1} e^{j\pi d m} \text{sinc } d_m \quad (\text{A2.1-35})$$

The solution to Equation (A2.1-35) is easily found from the solution to (A2.1-32) since the values of the  $X_n$  contributions for negative  $m$  are just the complex conjugates of contributions for positive  $m$ . The resulting value of  $X_n$  for  $M = 100$  and  $-1 \leq d \leq 1$  is

$$X_n \approx Ae^{j\theta} \times 0.22$$

This result assumes that the time required for  $d$  to sweep from  $-1$  to  $+1$  is just  $T$  seconds, the record length. The time required for the carrier to sweep from a value corresponding to  $d = -1$  to  $d = 1$  is:

$$T_{2d} = \frac{2B \times 2\pi}{\dot{W}} = \frac{4\pi B}{\dot{W}} \quad (\text{A2.1-36})$$

If  $T_{2d}$  is less than  $T$ , the average value of  $X_n$  is:

$$\bar{X}_n \approx 0.22 \frac{T_{2d}}{T} Ae^{j\theta} \quad (\text{A2.1-37})$$

$$= \frac{2.76}{\dot{W}} BAe^{j\theta}, \quad \dot{W} > \frac{4\pi B}{T} \quad (\text{A2.1-38})$$

Hence the amplitude of the DFT output is reduced to as little as 22% of the non-drifting case, and can be much less if the drift is quite rapid.



### Response of the DFT to Noise

Next we consider the effect of the DFT on noise. We assume that the input is Gaussian White noise with zero mean and power spectral density  $N_0$ . We assume that this noise is passed through an ideal bandpass filter of bandwidth  $B_T$ . The resulting bandpass noise can be written in the Rice representation (Viterbi, 1966) as:

$$n(t) = n_c(t) \cos W_c t + n_s(t) \sin W_c t \quad (A2.1-39)$$

where  $W_c$  is the center frequency of the bandpass filter, and  $n_c(t)$  and  $n_s(t)$  are low frequency independent Gaussian random processes. The variances of  $n(t)$ ,  $n_c(t)$  and  $n_s(t)$  all equal  $N_0 B_T$ , and we shall call this variance  $\sigma^2$ .

$$\sigma_n^2 = \sigma_{n_c}^2 = \sigma_{n_s}^2 = N_0 B_T \equiv \sigma^2 \quad (A2.1-40)$$

The quadrature detection process used in the receiver yields the complex low frequency (envelope):

$$x(t) = n_c(t) + j n_s(t) \quad (A2.1-41)$$

The  $k$ th sample of  $x(t)$  is

$$x_k = a_k + j b_k \quad (A2.1-42)$$

where  $a_k$  and  $b_k$  are defined respectively as the  $k$ th samples of  $n_c(t)$  and  $n_s(t)$ . Since the original processes  $n_c(t)$  and  $n_s(t)$  are Gaussian, the samples  $a_k$  and  $b_k$  are likewise Gaussian with zero mean and the same variance .

$$\sigma_a^2 = \sigma_b^2 = \sigma^2 \quad (A2.1-43)$$

The DFT of the complex samples  $x_k$  is:

$$X_n = \frac{1}{N} \sum_{k=0}^{N-1} x_k e^{-j2\pi nk/N}, \quad n = 0, 1, 2, \dots, N-1 \quad (A2.1-44)$$

$$= \frac{1}{N} \sum_{k=0}^{N-1} a_k e^{-j2\pi nk/N} + j \frac{1}{N} \sum_{k=0}^{N-1} b_k e^{-j2\pi nk/N} \quad (\text{A2.1-45})$$

$$\text{Since } e^{-j2\pi nk/N} = \cos \frac{2\pi nk}{N} - j \sin \frac{2\pi nk}{N} \quad (\text{A2.1-46})$$

Equation (A2.1-45) can be written as

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} (a_k \cos \frac{2\pi nk}{N} + b_k \sin \frac{2\pi nk}{N}) + j \frac{1}{N} \sum_{k=0}^{N-1} (b_k \cos \frac{2\pi nk}{N} - a_k \sin \frac{2\pi nk}{N}) \quad (\text{A2.1-47})$$

This result can be written as:

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} c_k + j \frac{1}{N} \sum_{k=0}^{N-1} d_k \quad (\text{A2.1-48})$$

where:

$$c_k = a_k \cos \frac{2\pi nk}{N} + b_k \sin \frac{2\pi nk}{N}$$

$$d_k = b_k \cos \frac{2\pi nk}{N} - a_k \sin \frac{2\pi nk}{N}$$

Clearly  $c_k$  and  $d_k$  are Gaussian random variables since the sum of Gaussian random variables is Gaussian (the cosine and sine terms are just constant weights for any given  $k$ ). Similarly the two sums in Equation (A2.1-48) are Gaussian, and hence  $x_n$  is a complex Gaussian random variable. It is easily shown that the mean values of  $c_k$  and  $d_k$  are zero. The variance of  $c_k$  is:

$$\begin{aligned} \sigma_{c_k}^2 &= E\{(a_k \cos^2 \frac{2\pi nk}{N} + b_k \sin^2 \frac{2\pi nk}{N})^2\} \\ &= E(a_k^2) \cos^2 \frac{2\pi nk}{N} + E(b_k^2) \sin^2 \frac{2\pi nk}{N} \\ &= \sigma^2 (\cos^2 \frac{2\pi nk}{N} + \sin^2 \frac{2\pi nk}{N}) \\ &= \sigma^2 \end{aligned} \quad (\text{A2.1-49})$$

Since the  $c_k$  are assumed to be independent over  $k$ , the variance of the sum of the  $c_k$  is the sum of the variances. Hence the variance of  $\frac{1}{N}$  times the first sum is  $\sigma^2$ . The same arguments apply to the second sum. In summary,  $X_n$  is a complex random variable whose real and imaginary terms are independent Gaussian with zero mean and variance  $\sigma^2$ . Equation (A2.1-48) can be written in polar form as:

$$X_n = R_n \angle \theta \quad (\text{A2.1-50})$$

We review next the well-known (Papoulis, 1965) distributions of the three random variables  $R_n$ ,  $\theta$ , and  $R_n^2 \equiv 2P$

The random variable  $R_n$  is Rayleigh,

$$f(R_n) = \frac{R_n}{\sigma^2} e^{-R_n^2/2\sigma^2}, \quad 0 \leq R_n < \infty, \quad (\text{A2.1-51})$$

The angle  $\theta$  is uniformly distributed.

$$f(\theta) = \frac{1}{2\pi}, \quad -\pi \leq \theta \leq \pi \quad (\text{A2.1-52})$$

The random variable  $P$ , which is the power of  $X_n$ , is exponentially (or Gamma) distributed:

$$f(P) = \frac{1}{\sigma^2} e^{-P/\sigma^2} \quad (\text{A2.1-53})$$

We now have the statistics of the output of the DFT when the input is Gaussian noise. Next we turn our attention to the statistics of the sum of a number of DFT outputs. This sum can be taken directly on the  $X_n$ , or on the power  $P$ , if a square law detector is first applied to the  $X_n$ . Consider the former case first.

We assume that the  $X_n$  being summed are independent complex Gaussian random variables. Let the sum of  $M$  values of  $X_n$  be called  $S$ .

$$S = \sum_{n=0}^{M-1} X_n \quad (A2.1-54)$$

The real and imaginary terms add but since these are independent and Gaussian, the two terms are both Gaussian, with mean zero and variance  $M\sigma^2$ . Hence  $S$  is still a complex Gaussian random variable. Its amplitude, angle and power are given by Equations (A2.1-51) - (A2.1-53) where  $\sigma^2$  is replaced by  $M\sigma^2$ .

We next consider the sum of  $M$  independent values of power  $P$ . We call this sum  $P_M$ .

$$P_M = \sum_{m=0}^{M-1} P_m \quad (A2.1-55)$$

Where the  $P_m$  are the powers which are to be added. They are exponentially distributed as given by Equation (A2.1-53). The mean and standard deviation of  $P_m$  both equal  $\sigma^2$ .

The sum of  $M$  independent exponentially distributed random variables is known to follow the Gamma distribution (Papoulis, 1965).

$$f(P_M) = \frac{P_M^{M-1}}{\sigma^{2M} (M-1)!} e^{-P_M/\sigma^2} \quad (A2.1-56)$$

where the mean and standard deviation of  $P_M$  are:

$$E[P_M] = M \sigma^2 \quad (A2.1-57)$$

$$\sigma_{P_M} = \sqrt{M} \sigma^2 \quad (A2.1-58)$$

The Gamma distribution becomes increasingly Gaussian-like as  $M$  increases, that is, as the number of summed terms grows. This is as one would expect from the Central Limit Theorem.

Note that if the  $P_m$  in Equation (A2.1-55) represent signal power, and hence are identical for each  $m$ , the total signal power is  $MP_m$ . That is, the signal power grows in proportion to  $M$ . On the other hand, the standard deviation of the total noise power is growing as  $\sqrt{M}$ , as Equation (A2.1-58) shows.

Hence the ratio of the total signal power to the standard deviation of the total noise power increases in proportion to  $\sqrt{M}$ , whereas the SNR increases in proportion to  $M$  if the summation is carried out on the amplitude instead of the power. Thus it is more desirable to carry out the summation over amplitude values if this is possible. Unfortunately, this cannot be done if the amplitudes are independent and random, or appear to be so because of the action of the DFT in the MCSA.

#### References

1. Papoulis, A., Probability, Random Variables, and Stochastic Processes, McGraw-Hill, New York, 1965.
2. Peebles, P., Communication System Principles, Addison-Wesley, Reading, Mass., 1976.
3. Viterbi, A.J., Principles of Coherent Communication, McGraw-Hill, 1966.

### Appendix A3.1

#### Polarization Synthesis

Given a desired type of polarization, as specified by the angles  $\delta$  and  $\gamma$ , what are the weighting factors needed to combine the outputs of two orthogonal antennas in the form

$$z = K_1 x + K_2 y$$

in such a way as to synthesize an antenna of that polarization?

In general,  $K_1$  and  $K_2$  are complex constants. Writing them in polar form, we have

$$Z = |K_1|e^{j\theta_1} x + |K_2|e^{j\theta_2} y$$

The value of  $\delta$  is the (specified) phase angle difference that a signal will have between the x and y antennas. To optimize reception, this phase difference must be cancelled, which gives the relation,

$$\delta + (\theta_1 - \theta_2) = 0$$

We are free to choose any values of  $\theta_2$  and  $\theta_1$ , which satisfy this equation, so let us choose for simplicity,

$$\theta_1 = 0 \text{ and } \theta_2 = \delta$$

The value of  $\gamma$  defines the relative magnitudes of the received signals, in the form

$$\tan \gamma = \frac{|K_2|}{|K_1|}$$

We are free to choose any values of  $K_1$  and  $K_2$  that satisfy this equation. Therefore let us also impose the condition that power be conserved by the combining operation, or in other words,

$$|K_1|^2 + |K_2|^2 = 1$$

Solving these two simultaneous equations for  $|K_1|$  and  $|K_2|$ , we obtain

$$|K_1| = \cos \gamma$$

$$|K_2| = \sin \gamma$$

Thus the required equation for combining the x and y outputs to match the polarization type ( $\delta, \gamma$ ) is

$$Z = x \cos \gamma + y e^{j\delta} \sin \gamma \quad (\text{A3.1-1})$$

From (3-1) and (3-2) we can obtain:

$$\cos \gamma = \frac{|x|}{\sqrt{|x|^2 + |y|^2}}$$

$$\sin \gamma = \frac{|y|}{\sqrt{|x|^2 + |y|^2}}$$

$$e^{j\delta} = e^{j(\angle y - \angle x)}$$

Then substituting these measured quantities into equation (A3.1-1), we obtain finally,

$$Z = \frac{|x|}{\sqrt{|x|^2 + |y|^2}} x + \frac{|y|}{\sqrt{|x|^2 + |y|^2}} y e^{j(\angle x - \angle y)} \quad (\text{A3.1-2})$$

If the outputs of the x and y antennas are combined according to this equation, the result will be the same as if a single antenna that was exactly matched to the polarization type of the incoming signal had been used, regardless of the polarization type of the signal.

#### Effects of Noise

Thus far, the effects of noise have been neglected, for the sake of simplicity. Noise causes the measured polarization type to have some probable error, and therefore the synthesized polarization type does not

exactly match that of the signal. Before discussing this effect explicitly, it is interesting to note that equation (A3.1-2) reduces to a nearly trivial equation in the complete absence of noise. If we write

$$\begin{aligned}x &= |x| e^{j\angle x} \\y &= |y| e^{j\angle y}\end{aligned}$$

Then,

$$\begin{aligned}Z &= \frac{1}{\sqrt{|x|^2 + |y|^2}} \left[ |x|^2 e^{j\angle x} + |y|^2 e^{j\angle y} e^{j(\angle x - \angle y)} \right] \\Z &= \frac{1}{\sqrt{|x|^2 + |y|^2}} \left[ e^{j\angle x} (|x|^2 + |y|^2) \right] \\Z &= \sqrt{|x|^2 + |y|^2} e^{j\angle x}\end{aligned}$$

The fact that the phase of  $Z$  is the same as the phase of  $x$  has no particular physical significance, and stems solely from the previous choice of  $\theta_1$  and  $\theta_2$ . The fact that the power in  $Z$  is the sum of the powers in  $x$  and  $y$  stems from the conservation of power constraint placed on  $|K_1|$  and  $|K_2|$  earlier. Thus we conclude that in the complete absence of noise, the polarization matching method extracts all of the available signal power, since

$$P_Z = Z^2 = x^2 + y^2 = P_x + P_y$$

Consider now the case when the signal to noise ratio (SNR) is small. The point  $M$  on the Poincare sphere in Figure 1 now marks the maximum value of a probability density function on the surface of the sphere. In order to increase the accuracy of the measured polarization type, it is necessary to average the weighting factors  $K_1$  and  $K_2$  over time.

C-4



The magnitudes of the factors then become

$$|K_1| = \left( \frac{|x|}{\sqrt{|x|^2 + |y|^2}} \right)$$

$$|K_2| = \left( \frac{|y|}{\sqrt{|x|^2 + |y|^2}} \right)$$

Note that it is not exactly correct to factor out the denominator of these expressions, or to regard it as the average power. To do so implies interchanging the order of averaging and computing, which leads to a somewhat different result.

An average phase difference between the x and y antennas must also be calculated. This could be done by simply averaging their phase difference over time. A statistically better average is the phase angle of the complex correlation coefficient between the two antenna outputs, given by

$$\mu_{xy} = \frac{\overline{xy^*}}{\sqrt{\overline{|x|^2}} \sqrt{\overline{|y|^2}}} \quad ; \quad (0 \leq |\mu_{xy}| \leq 1)$$

Since the denominator of  $\mu_{xy}$  is real, it does not affect its phase angle.

Thus

$$\angle \mu_{xy} = \angle \overline{xy^*} = \tan^{-1} \frac{\text{Im}(\overline{xy^*})}{\text{Re}(\overline{xy^*})}$$

This is a weighted average phase angle, where the weighting factors are the product of the amplitudes. This may be seen by considering the polar coordinate form of x and y.

$$\text{Let } x = r_x e^{j\alpha_x}$$

$$y = r_y e^{j\alpha_y}$$

Then

$$xy^* = r_x r_y e^{j(\alpha_x - \alpha_y)}$$

If we have N time samples of x and y, then

$$\overline{xy^*} = \frac{1}{N} \sum_{i=1}^N r_{xi} r_{yi} e^{j(\alpha_{xi} - \alpha_{yi})}$$

from which the weighted average effect may be seen. When the product  $r_{xi} r_{yi}$  is large, its corresponding phase difference will be weighted heavily. That product corresponds to cross power, so in fact instantaneous cross power is the weighting factor.

Summarizing these results, in the presence of noise we combine the outputs of the x and y antennas according to

$$Z = K_1 x + K_2 y \quad (\text{A3.1-3})$$

where

$$K_1 = \left( \frac{|x|^2}{|x|^2 + |y|^2} \right)$$

$$K_2 = \left( \frac{|y|^2}{|x|^2 + |y|^2} \right) e^{j\delta}$$

$$\delta = \tan^{-1} \frac{\text{Im}(\overline{xy^*})}{\text{Re}(\overline{xy^*})}$$

The factors  $K_1$  and  $K_2$  emphasize the antenna that has the strongest signal, and hence the best SNR. Since the time averaged information must be used to combine even the first signals, a time delay must occur before the combined output is available. This may be accomplished by recording the signal, or by appropriate data processing. However, it will turn out that the final expression for generalized coherence to be derived below may be computed continuously in quasi-real time, using "pipeline" techniques similar to those used in real-time FFT calculations.

The above results may be generalized to any two orthogonal antennas. The fact that x and y oriented linear antennas were being used only affects

the definition of the polarization angles  $\gamma$  and  $\delta$  on the Poincare sphere. Horizontal linear polarization is chosen by convention to be the origin of the coordinate system. However, one could with equal validity choose any other point (such as left-circular) as the origin, and then define new polarization angles exactly analogous to  $\delta$  and  $\gamma$ , relating the amplitudes and phase angles of the two new antennas. The angles  $\delta$  and  $\gamma$  are only intermediate variables, and once they are eliminated by substituting their values in terms of the observable quantities into equation (A3.1-3), their meaning and definition become irrelevant. Thus equation (A3.1-3) is true for any two orthogonally polarized antennas. For example, if left and right hand circularly polarized antennas were used,  $x$  and  $y$  could simply be replaced by  $R$  and  $L$ .

## Appendix A3.2

### Coherency Matrix and Degree of Polarization

A matrix known as the Coherency Matrix contains all of the polarization information about a signal (Ko, 1962 and Beran et al., 1964). It is

$$[c] = \begin{bmatrix} \overline{xx^*} & \overline{xy^*} \\ \overline{x^*y} & \overline{yy^*} \end{bmatrix}$$

where  $x$  and  $y$  are the complex output signals from two orthogonal receiving antennas. The total signal power is

$$P_s = \text{trace } [c] = \overline{xx^*} + \overline{yy^*}$$

The degree of polarization is

$$\begin{aligned} m &= \sqrt{1 - \frac{4 \det [c]}{(\text{trace } [c])^2}} \\ &= \sqrt{1 - \frac{4(\overline{xx^*} \overline{yy^*} - \overline{xy^*} \overline{x^*y})}{(\overline{xx^*} + \overline{yy^*})^2}} \end{aligned}$$

Using the fact that  $\overline{xx^*} = P_x$ ,  $\overline{yy^*} = P_y$ , and  $\overline{xy^*} \overline{x^*y} = P_{xy}$  we have

$$m = \sqrt{1 - \frac{4(P_x P_y - P_{xy}^2)}{(P_x + P_y)^2}}$$

This may also be rearranged to

$$m = \frac{1}{P_x + P_y} \sqrt{(P_x - P_y)^2 + 4P_{xy}^2}$$

### Appendix A3.3

#### Effects of Noise on the Measured Degree of Polarization

The measured degree of polarization is affected by noise in two distinct ways. The first is that it has statistical fluctuations, and therefore must be computed using coherence matrix values that are averaged over a significant length of time.

Degree of polarization behaves like a correlation function in that if only one time sample is used, its measured value will always have a value of 1. As more and more time samples are used, its measured value will approach the true value, from above (Bendat, 1958; Lee, 1970; Ottes, et al, 1978). This behavior is illustrated in Figures A3.3.2 and A3.3.3 for various signal-to-noise ratios.

The probability density for the degree of polarization is shown in Figure A3.3.1. These curves were computed by adding signals to random noise in such a way as to simulate the actual operation of experimentally measuring the degree of polarization. The curves apply to a single antenna which is matched to the signal polarization. Each curve is a smoothed histogram of 5000 total samples. The integration time used for these curves is 25 data points.

The second effect of noise on the measured degree of polarization is that receiver noise cannot be directly distinguished from unpolarized signals. A strong partially polarized signal may have the same measured degree of polarization as a weak polarized signal. This effect manifests itself as a measured degree of polarization that is always lower than the

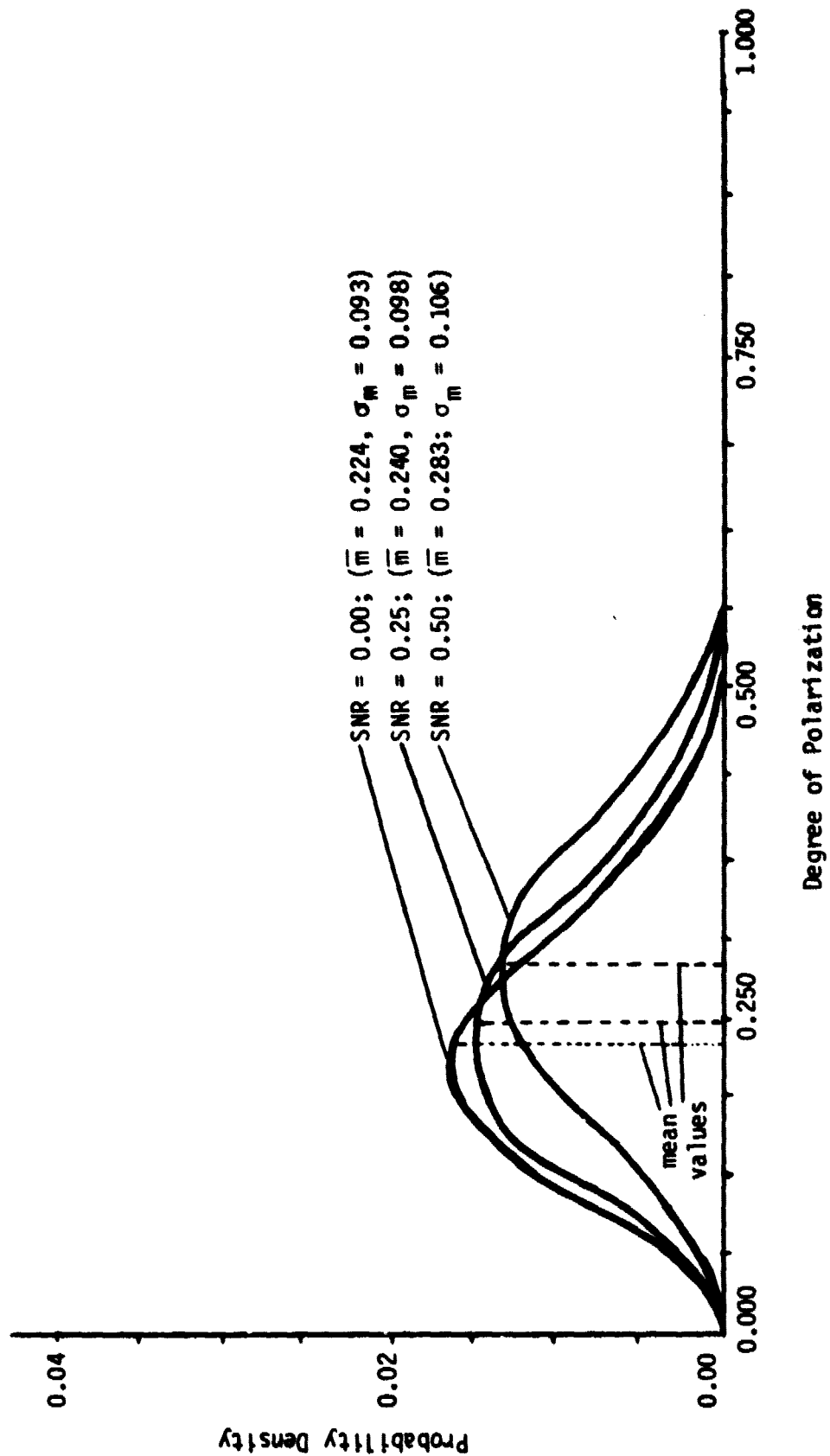
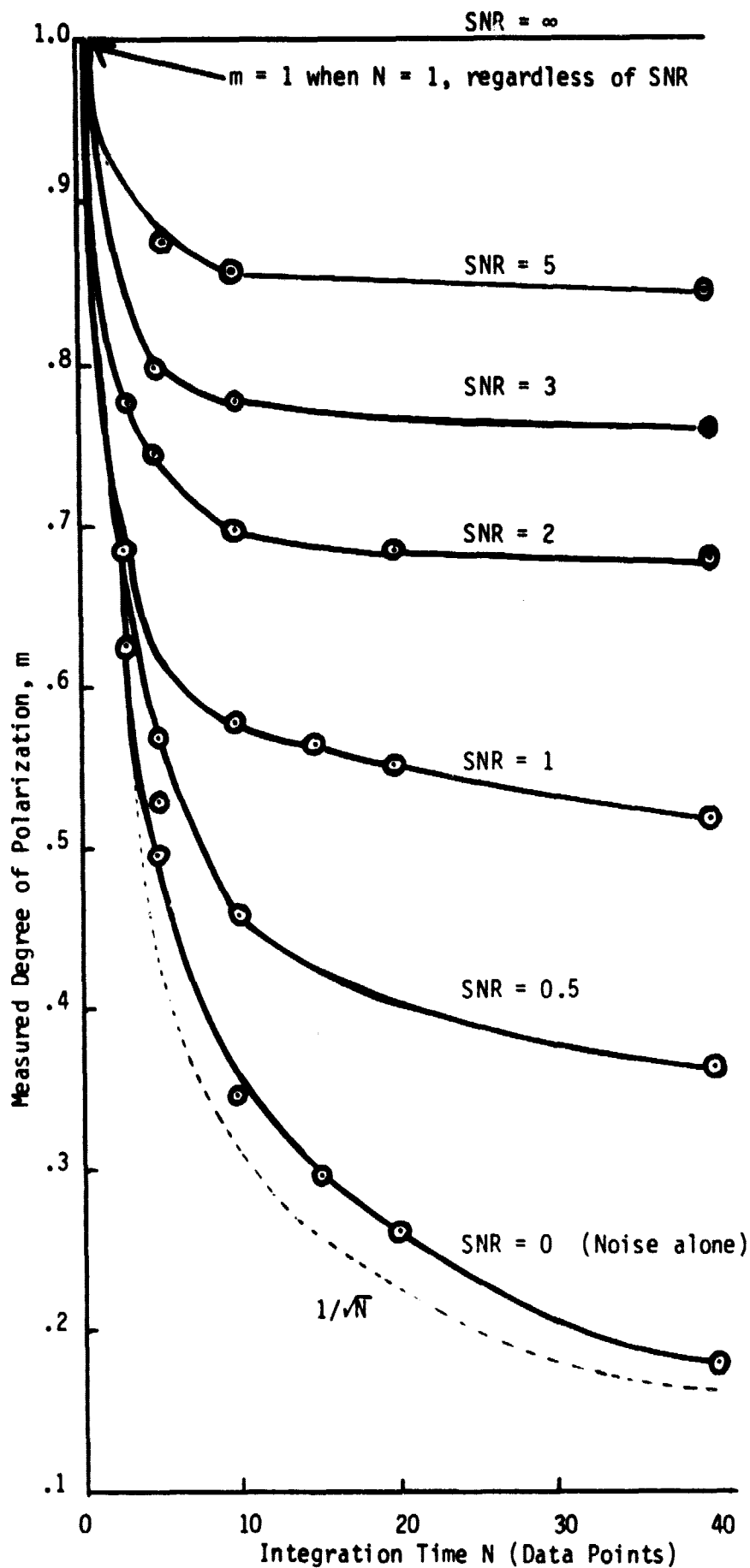


Figure A3.3.1 - Computed Probability Density for Degree of Polarization



All approach  

$$\frac{\text{SNR}}{1 + \text{SNR}}$$
  
 as  $N \rightarrow \infty$

Figure A3.3.2 - Measured Degree of Polarization, when the True Degree of Polarization is 1.0.

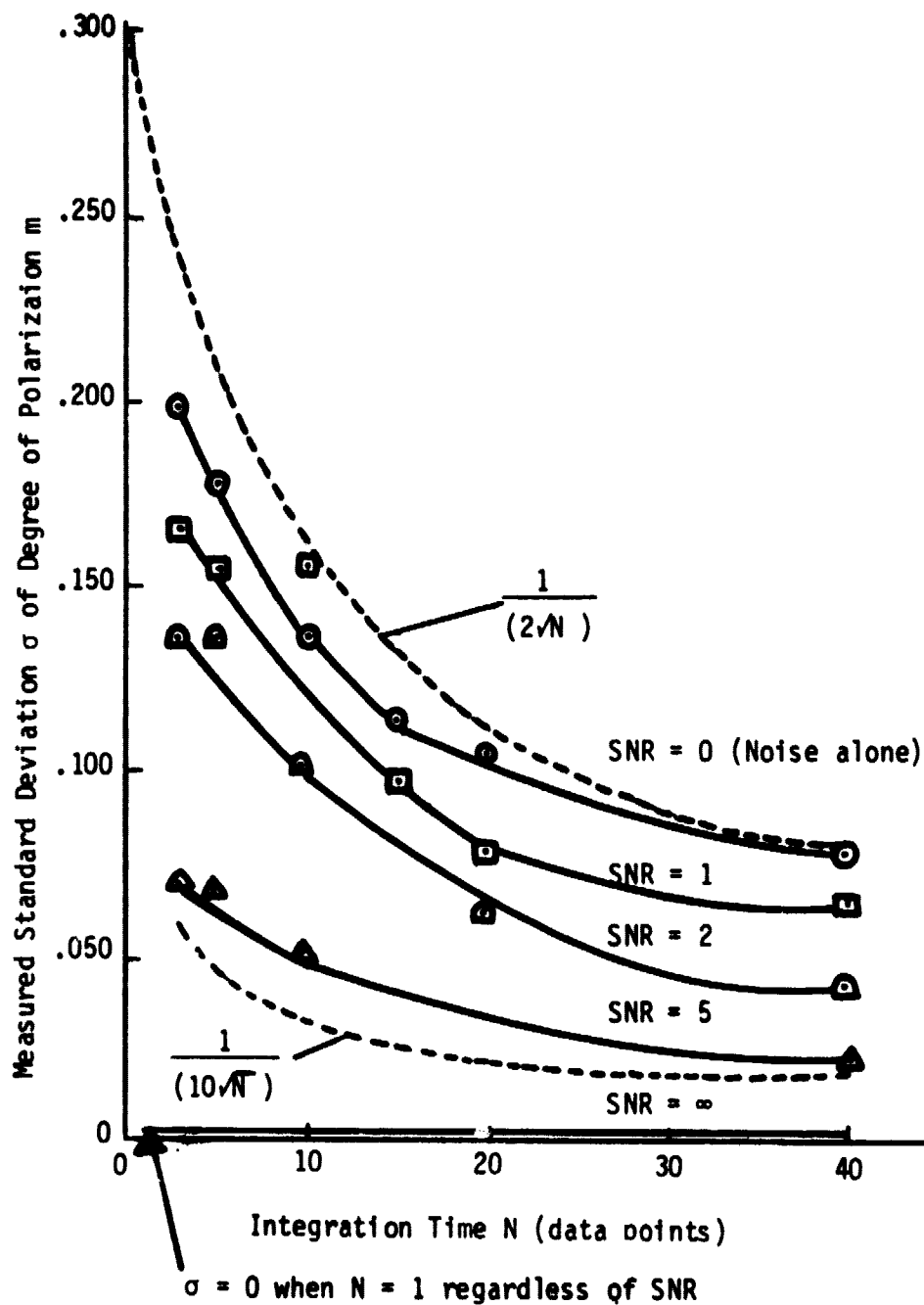


Figure A3.3.3 - Measured Standard Deviation of Degree of Polarization, When the True Degree of Polarization is 1.0.



true degree of polarization. When the SNR is large, the measured and true values are the same, and when the SNR is very small, the measured value is nearly zero.

This effect may be quantified as follows: The received signal powers in the two antennas are

$$P_x = P_s \left( \frac{1-m}{2} + m \cos^2 \frac{\overline{MM}^T}{2} \right)$$

$$P_y = P_s \left( \frac{1-m}{2} + m \cos^2 \left( -\frac{\pi - \overline{MM}^T}{2} \right) \right)$$

The total received signal power is, after simplification

$$P_r = P_x + P_y = P_s$$

To this we add the receiver noise, in the form

$$P_n = \frac{P_s}{\text{SNR}}$$

Since the receiver noise power in each receiver is the same, we add half of the total to each. After the polarization type has been matched, the  $\cos^2$  factor is 1, so the total received power in the polarization-matched antenna is

$$P_r = P_s \left( \frac{1-m}{2} + \frac{1}{2\text{SNR}} + m \right)$$

If we lump together the two noise-like components of the power (receiver noise and unpolarized signal), and regard them as being caused by some other unpolarized signal, then the original unpolarized power of that signal must have been twice that measured (since any antenna receives only half the unpolarized power). Thus

$$P_{\text{unpolarized}} = P_s \left( 1 - m + \frac{1}{\text{SNR}} \right)$$

The polarized component is  $mP_s$ , so we can compute the measured degree of polarization  $m'$ , as

$$m' = \frac{\text{polarized power in original signal}}{\text{total power in original signal}}$$

$$m' = \frac{mP_s}{P_s (1 - m + \frac{1}{\text{SNR}} + m)}$$

$$m' = m \left( \frac{\text{SNR}}{1 + \text{SNR}} \right)$$

This effect is illustrated in Figure A3.3.2, as the asymptotic values approached as  $N \rightarrow \infty$ . If it were desired to measure the true degree of polarization accurately, one could measure the SNR separately (by comparing the total measured power at the time when the signal is sought, with a globally averaged total power), and then invert the above relation,

$$m = m' \left( \frac{1 + \text{SNR}}{\text{SNR}} \right)$$

This does not become infinite for small SNR, because  $m'$  is also small in that case. In the limit both SNR and  $m'$  become zero simultaneously, leaving  $m$  indeterminate. This just reflects the physical reality that you cannot measure the polarization of a signal that is not there.

Although in principle one could correct the measured degree of polarization for both the finite-sample effect and the receiver noise effect, it is not necessary to do so for the application at hand. First it should be pointed out that the two effects are in opposite directions, so that they tend to partially cancel. Second, the measured degree of polarization is always a monotonically increasing function of the true degree of polarization, so that a thresholding operation may always be carried out successfully, even in the presence of the two above-mentioned effects.

#### Appendix A3.4

##### Linear Phase Invariance Property of Coherent Power

To demonstrate the linear phase invariance property, let the phase of the signal  $z$  advance linearly by an amount  $\Delta\theta$  radians per second. Then if we explicitly write out the averaging process for coherent power,

$$P_c = \left| \frac{1}{N-1} \sum_{i=1}^{N-1} z_{i+1} z_i^* \right|$$

and change the  $z$ 's to polar coordinates,

$$z_i = A_i e^{j(\theta + i\Delta\theta)}$$

where  $\theta$  is the original (arbitrary) phase angle of  $z$ .

Then

$$P_c = \left| \frac{1}{N-1} \sum_{i=1}^{N-1} A_i e^{j(\theta + i\Delta\theta)} A_{i+1} e^{-j(\theta + (i+1)\Delta\theta)} \right|$$

$$P_c = \left| \frac{1}{N-1} \sum_{i=1}^{N-1} A_i A_{i+1} e^{-j\Delta\theta} \right|$$

$$P_c = \left| \frac{1}{N-1} \right| \cdot \left| e^{-j\Delta\theta} \right| \cdot \left| \sum_{i=1}^{N-1} A_i A_{i+1} \right|$$

$$P_c = \frac{1}{N-1} \left| \sum_{i=1}^{N-1} A_i A_{i+1} \right|$$

Since this is independent of  $\Delta\theta$ , we see that the linear phase shift has no effect on the outcome. Note, however, that any other phase shifts (non-linear or random) would cause  $P_c$  to become smaller, thereby illustrating its sensitivity to phase coherence.

### Appendix A3.5

#### Generalized Coherence Mathematical Details

From Equations (3-5), (3-9), and (3-10), we obtain

$$GCV = \frac{2m}{1+m} \left| \frac{1}{N-1} \sum_{i=1}^{N-1} z_{i+1} z_i^* \right| \quad (A3.5-1)$$

where

$$m = \frac{1}{p_x + p_y} \sqrt{(p_x - p_y)^2 + 4p_{xy}}$$

$$z_i = K_1 x_i + K_2 y_i$$

and  $K_1$  and  $K_2$  were defined earlier in equation (3-5). Substituting for the  $z_i$  values, the summation term becomes

$$\begin{aligned} \text{Summation} &= \sum_{i=1}^{N-1} (K_1 x_{i+1} + K_2 y_{i+1}) (K_1^* x_i^* + K_2^* y_i^*) \\ &= \sum_{i=1}^{N-1} (|K_1|^2 x_{i+1} x_i^* + K_2 K_1^* y_{i+1} x_i^* + K_1 K_2^* x_{i+1} y_i^* \\ &\quad + |K_2|^2 y_{i+1} y_i^*) \end{aligned} \quad (A3.5-2)$$

From equation (3-5), we obtain

$$|K_1|^2 = \left[ \frac{1}{N} \sum_{i=1}^N \frac{|x_i|}{\sqrt{|x_i|^2 + |y_i|^2}} \right]^2$$

which may be regarded conceptually as the power in the x antenna, normalized to the total power; call it  $P_{xn}$ .

$$|K_2|^2 = \left[ \frac{1}{N} \sum_{i=1}^N \frac{|y_i|}{\sqrt{|x_i|^2 + |y_i|^2}} \right]^2$$

which may be regarded conceptually as the power in the y antenna, normalized to the total power. Call it  $P_{yn}$ .

$$K_2 K_1^* = \left[ \frac{1}{N} \sum_{i=1}^N \frac{|y_i|}{\sqrt{|x_i|^2 + |y_i|^2}} \right] \left[ \frac{1}{N} \sum_{i=1}^N \frac{|x_i|}{\sqrt{|x_i|^2 + |y_i|^2}} \right] e^{j\delta}$$

which may be regarded conceptually as the cross power between the x and y antennas, normalized to the total power, with a phase lag of  $\delta$ . Call it  $P_{xyn} e^{-j\delta}$ .

$$K_1 K_2^* = \left[ \frac{1}{N} \sum_{i=1}^N \frac{|x_i|}{\sqrt{|x_i|^2 + |y_i|^2}} \right] \left[ \frac{1}{N} \sum_{i=1}^N \frac{|y_i|}{\sqrt{|x_i|^2 + |y_i|^2}} e^{-j\delta} \right]$$

which may be regarded conceptually as the ratio of the cross power between the x and y antennas, to the total power, with a phase lag of  $\delta$ . Call it  $P_{xyn} e^{-j\delta}$ . Then the summation becomes

$$\begin{aligned} \text{Summation} = \sum_{i=1}^N & \left[ P_{xn} x_{i+1} x_i^* + P_{xyn} e^{j\delta} y_{i+1} x_i^* \right. \\ & \left. + P_{xyn} e^{-j\delta} x_{i+1} y_i^* + P_{yn} y_{i+1} y_i^* \right] \quad (\text{A3.5-3}) \end{aligned}$$

The first term within the summation is the autocorrelation function of the signal in the x antenna, weighted by the normalized power in the x antenna. The last term within the summation is the autocorrelation function of the signal in the y antenna, weighted by the normalized power in the y antenna. The middle two terms within the summation are time-shifted cross-correlation functions between the signals in the x and y antennas, weighted by the normalized cross power in the x and y antennas. The two middle terms differ in that one has the time lag in x, and the other has it in y, so that both are included in a symmetrical way. The two middle terms also differ in their phases. Due to the polarization type of the signal, the phase of x is expected

to lead the phase of  $y$  by the angle  $\delta$ . This phase difference will on the average be cancelled by the  $e^{-j\delta}$  factor for  $y$ , and by the  $e^{j\delta}$  factor for  $y^*$ . The extent to which this phase difference cancellation occurs is a measure of the phase coherence of the signal both in time and between the two antennas.

Thus we see that the terms within the summation of equation (3-10) represent all of the possible time-shifted correlation functions of the signal, with each correlation function being weighted by its importance relative to the particular signal being received.

The multiplicative factor outside of the summation of equation (3-10) takes into account all of the possible non-time-shifted correlation functions of the signal, namely

$$\begin{aligned} P_x &= \frac{1}{N} \sum_{i=1}^N x_i x_i^* \\ P_y &= \frac{1}{N} \sum_{i=1}^N y_i y_i^* \\ P_{xy} &= \frac{1}{N} \sum_{i=1}^N x_i y_i^* \end{aligned}$$

In summary, the generalized coherence expression involves all the possible correlation functions of the signal. The separate effects of polarization and time coherence lose their individual identities, and thus generalized coherence may be regarded as a fundamental characteristic of the signal itself, apart from the way in which it is received and measured.

The correlation functions in equation (3-12) are merely generalizations of those found in the coherency matrix discussed earlier. This leads to

the suggestion of a Generalized Coherency matrix,

$$[GC] = \begin{bmatrix} \overline{\lambda x^*(\tau)} & \overline{yy^*(\tau)} \\ \overline{x^*y(\tau)} & \overline{yy^*(\tau)} \end{bmatrix}$$

where  $\tau$  is the time shift of the correlation function. When  $\tau = 0$ , this matrix reduces to the coherency matrix, which as stated previously completely defines the polarization of a signal. When  $\tau \neq 0$ , this matrix completely defines the time coherence properties of a signal. Thus the Generalized Coherency matrix completely defines both the polarization and time coherence properties of a signal.

## Appendix A3.6

### Generalized Coherence Statistics

Test cases showed that GCV is insensitive to the distribution of signal power between the two feed antennas, so all simulations added the signal to only one antenna, and the signal-to-noise ratio is therefore defined to be that in a polarization-matched antenna. The actual operational SNR is always lower than that, since the other antenna contributes an equal amount of noise and no additional signal. It is difficult to define a SNR which takes this polarization effect into account satisfactorily. If one simply takes the ratio of total signal power to total noise power (adding powers arithmetically), one obtains a SNR that is one half of that in a single matched antenna. On the other hand if one were to add together the output power signals from the two antennas, using simple square law detectors, the SNR would be  $\sqrt{2}$  less than that in a single matched antenna. The definition used here always underestimates the performance of GCV relative to total power detection.

In Figure 3.5 the open circles are the simulation outputs of 25 seconds integration on the sum of the powers of the two feed antennas. The solid line is the theoretical locus of these points, representing a SNR improvement of five, due the integration over 25 data points (i.e.,  $\sqrt{25} = 5$ ). The plus signs are the simulation outputs of GCV, computed over the same 25 data points. The dashed line is simply a smooth curve drawn through these points.



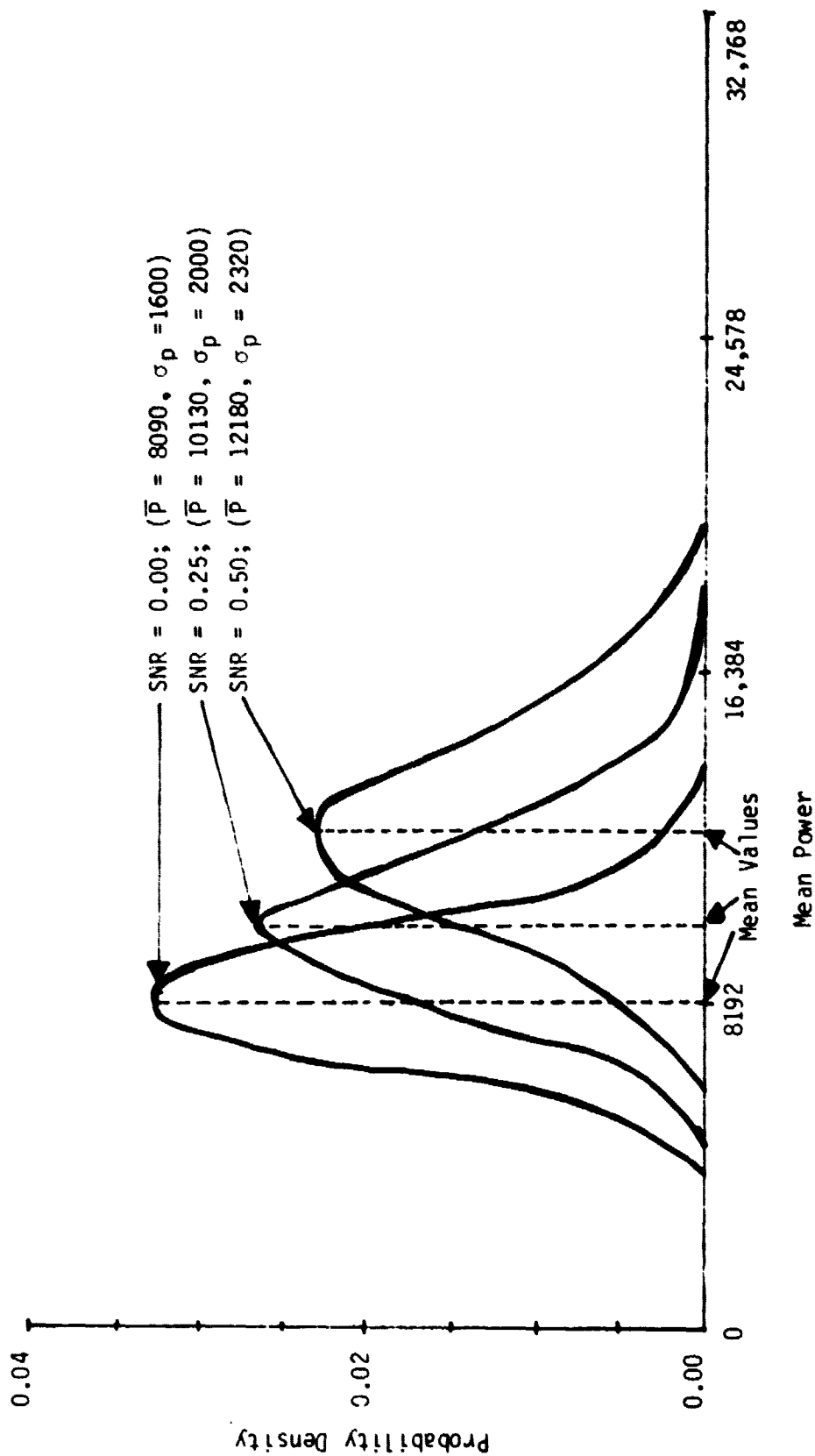


Figure A3.6.1 - Computed Probability Density for Time Integration of Total Power

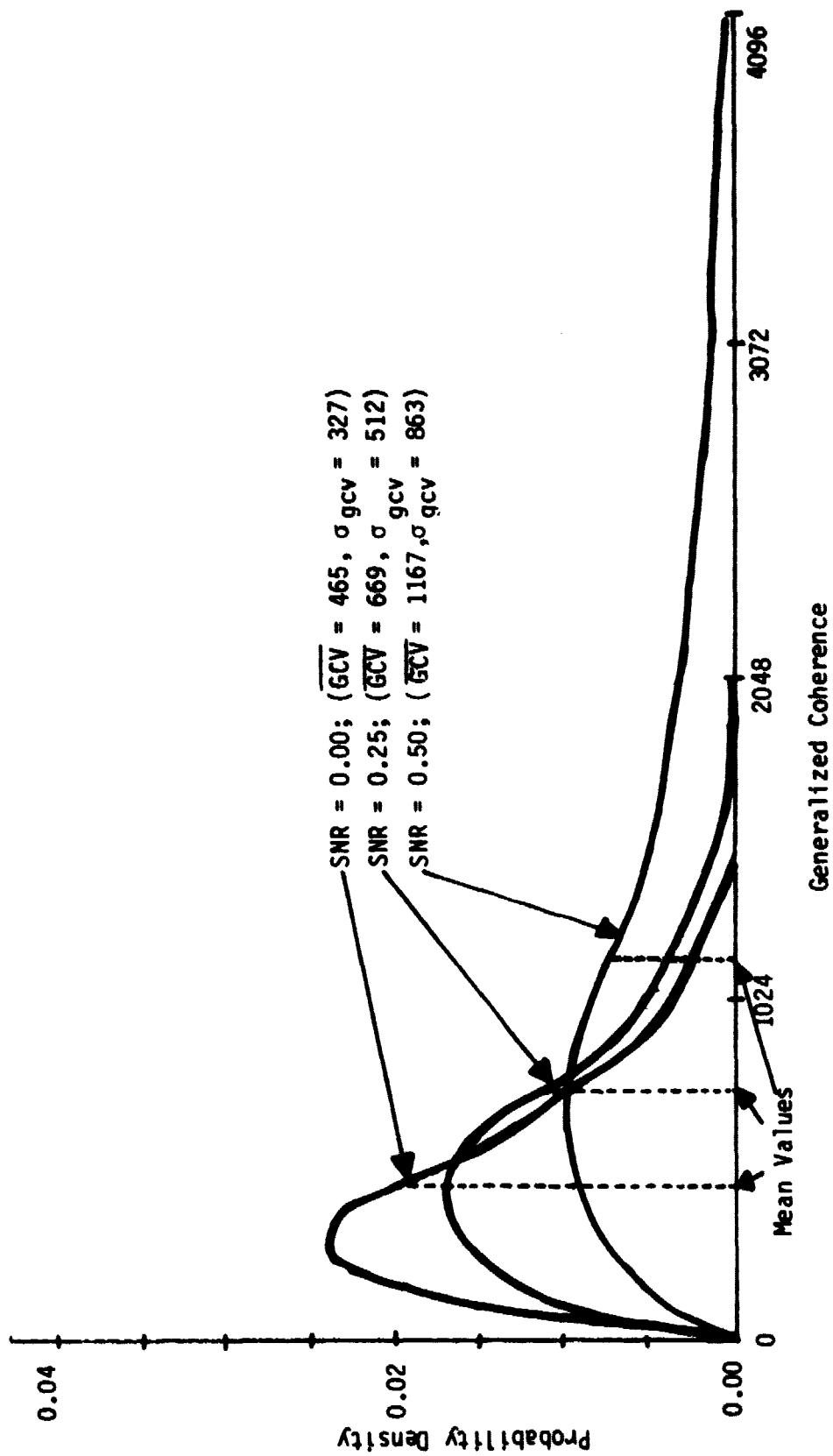


Figure A3.6.2 - Computed Probability Density for Generalized Coherence

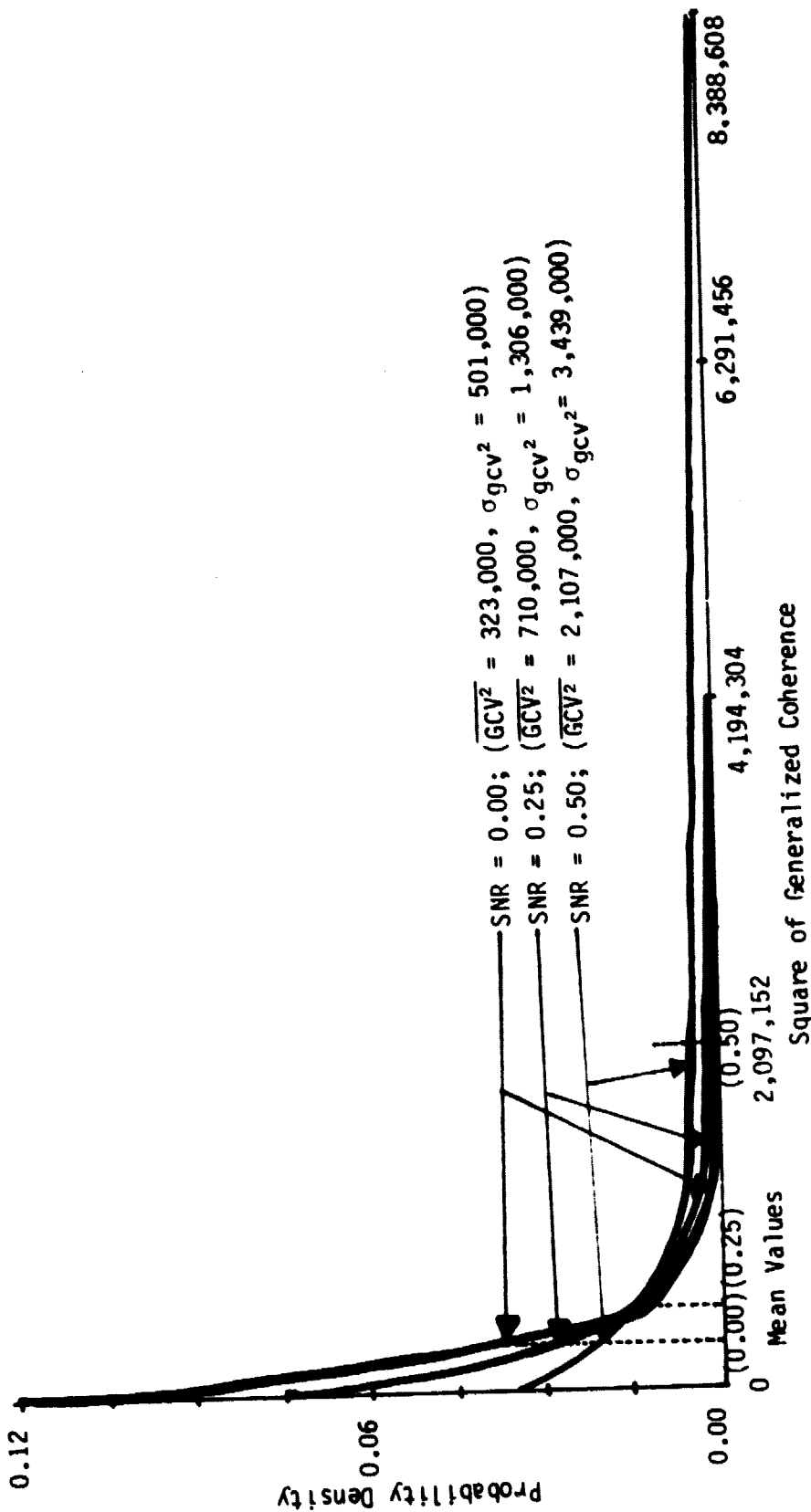


Figure A3.6.3 - Computed Probability Density for the Square of Generalized Coherence

For low SNR, additional data points were obtained for Figure 3.5 from the simulation probability density curves shown in Figures A3.6.1, A3.6.2, and A3.6.3. The points were obtained using a relation:

$$\text{OUTPUT SNR} = \frac{(\text{MEAN VALUE, AT DESIRED SNR}) - (\text{MEAN VALUE, AT ZERO SNR})}{\text{STANDARD DEVIATION, AT ZERO SNR}}$$

The power points are shown as solid circles, the GCV points as letter X's, and the square of GCV as solid triangles. Since the simulations for probability density involved a much greater number of total data points (5000) than did the simple comparison simulations (25), the experimental error is much smaller in the former case. This explains why the points obtained from the power probability density curves fall exactly on the theoretical curve, whereas the others exhibit a scatter.

### Appendix A3.7

#### The Possibility of Truly Coherent Detection

Generalized coherence does not require knowledge of the exact signal frequency within an MCSA output bin, because of its phase invariance property. The one remaining piece of information that is not used by generalized coherence is the phase angle of the degree of coherence (i.e., the phase angle of the complex autocorrelation function). This phase angle is in fact the average phase change per data point (i.e., the linear phase term). The offset frequency from the center of the bin is  $\frac{\angle \mu_{zz}}{\pi}$  (since  $f = \frac{d\theta}{dt}$ ), in units of binwidths. In principle, one could compute  $\angle \mu_{zz}$  after taking the polarization information into account, and then use it to apply a phase correction to the signal that compensates for its frequency within the MCSA bin. The procedure could be iterated if necessary, by computing a new  $\angle \mu_{zz}$  after the first corrections had been applied, and repeating the process. This would presumably permit a truly coherent detector to be used, in place of that portion of the generalized coherence that calculates  $|\sum z_{i+1} z_i^*|$ . Instead we would compute  $|\mu_{zz}| \left| \sum_{i=1}^N z_i \right|^2$  where  $|\mu_{zz}|$  now serves as a measure of the coherence of the signal.

Appendix A3.8  
Theory of ANOVA

Analysis of variance is deeply rooted in the history of statistical methodology (Cochrane and Cox, 1962, and Scheffe, 1950) and it is a prime tool for the handling of data obtained from designed experiments. As its name implies, it involves the systematic partitioning of the total variation in specially obtained data sets. The partitioning is done in order to assess the contribution each designated source of variability has on the total variance.

The form of analysis of variance that is appropriate for the SETI data matrix is the two-way ANOVA with cross classification. The columns of the data matrix represent the FREQUENCY in Hz, while the rows represent the TIME in seconds. Cross classification is obtained by having at least two observations per matrix cell. This "pseudo" set of replicated observations (Sec. 3.5, ANOVA, Performance) is necessary for the calculation of an interaction effect, and an interaction effect is necessary for the detection of certain types of signal. So, to consider the most general class of signal types, an  $r \times c$  matrix of time vs. frequency, with  $n$  observations per cell will be the format for the analysis of variance done on the data.

Let  $Y_{ijk}$  represent the observed response of the  $k^{\text{th}}$  replicate to the  $i^{\text{th}}$  level of the row treatments and the  $j^{\text{th}}$  level of the column treatments. Stated in a more physical fashion,  $Y_{ijk}$  represents the calculated (observed) polarized power of the  $k^{\text{th}}$  replicate at the  $i^{\text{th}}$  time period and the  $j^{\text{th}}$  frequency.

Each observed response,  $Y_{ijk}$  can be represented by the following model:

$$Y_{ijk} = \mu + \rho_i + \gamma_j + (\rho\gamma)_{ij} + \epsilon_{ijk} \quad (A3.8-1)$$

where  $i = 1, \dots, r$  and  $r$  is the number of levels of the first factor (time);

$j = 1, \dots, c$  and  $c$  is the number of levels of the second factor (frequency);

$k = 1, \dots, n$  and  $n$  is the number of replications for each row x column combination;

while  $\mu$  is the overall mean response, which, for this problem, is the average power.  $\rho_i$  represents the effect of the  $i^{\text{th}}$  level of the first factor averaged over  $c$  levels of the second factor; (the  $i^{\text{th}}$  time interval response values averaged over all frequencies).

$\gamma_j$  represents the effect of the  $j^{\text{th}}$  level of the second factor averaged over the  $r$  levels of the first factor (the  $j^{\text{th}}$  frequency averaged over time).

$(\rho\gamma)_{ij}$  represents the interaction between the  $i^{\text{th}}$  level of the first factor and the  $j^{\text{th}}$  level of the second factor (the interaction between the  $i^{\text{th}}$  time (second) and the  $j^{\text{th}}$  frequency).

$\epsilon_{ijk}$  represents the deviation of  $Y_{ijk}$  from the population mean response for the  $(ij)^{\text{th}}$  population.

To demonstrate how the total sum of squares can be partitioned into the sum of squares attributed to row and column and interaction effects, the following mathematical discussion is presented.

Consider the deviation obtained by subtracting the mean  $\mu$  in (A3.8-1) from  $Y_{ijk}$ , and substituting the following point estimates for the parameters:

$$\mu = \bar{Y}_{...}^*$$

$$\rho_i = \bar{Y}_{i...} - \bar{Y}_{...}$$

$$\phi_j = \bar{Y}_{.j.} - \bar{Y}_{...}$$

$$\hat{\epsilon}_{ijk} = Y_{ijk} - \bar{Y}_{ij.}$$

The following expression for the partition of the total deviation about the mean:

$$\begin{aligned} (Y_{ijk} - \bar{Y}_{...}) &= (\bar{Y}_{i..} - \bar{Y}_{...}) + (\bar{Y}_{.j.} - \bar{Y}_{...}) \\ &\quad + (\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...}) + (Y_{ijk} - \bar{Y}_{ij.}) \quad (A3.8-2) \end{aligned}$$

is obtained.

Three tests of hypothesis can be realized by looking at the computed F-ratios for rows, columns and interactions. For a given  $\alpha$ -level (a prespecified risk) the critical value can be found for the appropriate F-distribution. The null hypothesis of no-signal present is rejected if the computed F-ratio for rows, or columns or interactions exceeds its critical F-value.

The choice of  $\alpha$ , which is the risk allowed in rejecting the null hypothesis,  $H_0$ , even though it be true, can be made as small as desired

---

\* The dot (·) indicates summation over all values of the index replaced by the dot; i.e.,

$$\bar{Y}_{...} = \frac{1}{n \cdot r \cdot c} \sum_{i=1}^r \sum_{j=1}^c \sum_{k=1}^n Y_{ijk}$$



SOURCE OF VARIATION	DEGREES OF FREEDOM	SUM OF SQUARES	MEAN SUM OF SQUARES	F-RATIO
BETWEEN ROWS (time)	$r-1$	$SSR = nc \sum_{i=1}^r (\bar{Y}_{i..} - \bar{Y}_{...})^2$	$MSR = \frac{SSR}{r-1}$	$F_R = \frac{MSR}{MSE}$
BETWEEN COLUMNS (frequency)	$c-1$	$SSC = nr \sum_{j=1}^c (\bar{Y}_{.j.} - \bar{Y}_{...})^2$	$MSC = \frac{SSC}{c-1}$	$F_C = \frac{MSC}{MSE}$
INTERACTION (time x frequency)	$(r-1)(c-1)$	$SSI = n \sum_{i=1}^r \sum_{j=1}^c (\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...})^2$	$MSI = \frac{SSI}{(r-1)(c-1)}$	$F_I = \frac{MSI}{MSE}$
RESIDUAL	$rc(n-1)$	$SSE = \sum_{i=1}^r \sum_{j=1}^c \sum_{k=1}^n (\bar{Y}_{ijk} - \bar{Y}_{ij.})^2$	$MSE = \frac{SSE}{r \cdot c}$	
TOTAL	$r \cdot c \cdot n - 1$	$TSS = \sum_{i=1}^r \sum_{j=1}^c \sum_{k=1}^n (\bar{Y}_{ijk} - \bar{Y}_{...})^2$		

Table A3.8.1 - Two-Way Analysis of Variance with Cross Classification

(Type I error) However, a very small  $\alpha$  will result in a large Type II error, called  $\beta$ .  $\beta$  is the probability of accepting the null hypothesis when it is false. The relationship between Type I and Type II errors is indicated in Table A3.8.2.

		<u>DECISION</u>	
		ACCEPT	REJECT
<u>HYPOTHESIS</u>	$H_0$	NO ERROR	$\alpha$ TYPE I ERROR
	$H_1$	$\beta$ TYPE II ERROR	NO ERROR

Table A3.8.2 - Type I and Type II Errors

For the detection problem,  $\alpha$  represents the probability of affirming the presence of a signal when none is present; i.e., a FALSE ALARM.  $\beta$  represents the probability of MISSING A SIGNAL when one is present. A trade off between these two types of errors is always present in any statistical hypothesis testing. A decrease in one type of error results in an increase in the other type of error, for a fixed number of observations. Usually,  $\alpha$  is preset at some tolerable level, and the value of  $\beta$  is determined for various alternative hypotheses. The complement of  $\beta$ , namely  $(1 - \beta)$  is called the power of the test. Plots have been determined (Chemical Rubber Co., 1968) showing the power of ANOVA for given  $\alpha$  against various alternative hypotheses. The alternative hypotheses state that a signal is present. Obviously, the stronger the signal the higher the value of  $(1 - \beta)$ , and thus the greater the likelihood of its being detected.

Appendix A3.9  
ANOVA Performance

Experimental Conditions

1. All experiments were carried out at nine signal-to-noise ratio (SNR) levels, starting with 9 db and decreasing to -12 db in decrements of 3 db. The SNR is represented in the figures by the vertical axis. The horizontal axis represents the row, column, or interaction variance ratio depending upon the type of signal being analyzed.
2. Four types of signals were consistently analyzed. They are:
  - (a) non-drifting narrow-band signals,
  - (b) broad band pulsed signals,
  - (c) slowly drifting narrow-band signals, and
  - (d) drifting narrow band signals.

These four types represent probable forms of a meaningful ETI signal, and are diverse enough to trigger one (or more) of each type of hypothesis provided for by ANOVA.

3. Three factors that affect the quality of detection of signals by ANOVA were examined for each of the four signal types of all specified SNR. These factors are:
  - (a) width of signal
  - (b) frequency averaging
  - (c) number of bits used to represent the signal power (quantization level)
4. The data for all combinations of the three above-mentioned factors were each analyzed six times with six different noise fields, and the averages of the six results are shown in Figures A3.9. to A3.9. .

### Discussion of Variance Ratio vs. F-Ratio

The use of the expression "variance ratio" instead of the familiar F-ratio is in recognition of the fact that the distribution of the noise power data is exponentially distributed rather than Gaussian, hence the ratio of the squares of these random variables follows some distribution that is not yet known.\* Until the theoretical distribution is established, critical values appropriate to the F-distribution for various  $\alpha$ -levels can serve as conservative thresholds for the hypothesis tests. Empirical evidence of the distribution of the variance ratio for the noise data, seems to support this assumption.

### Description of Figures

Figure A3.9.1 depicts the representation of four types of signals in the noise field matrix.

Figures A3.9.2(a-d) demonstrate the sensitivity of ANOVA as a function of signal width. Widths of 1, 2 and 4 Hz bandwidth are superimposed upon the noise field according to the signal type pattern. As expected, an increase in signal width in general improves the detectability as is evidenced by the larger values of the variance ratio. In particular, Figure A3.9.2(a) depicts the response of ANOVA (as represented by the column variance ratio) to a non-drifting narrow-band signal. If a variance ratio of 1.3 represents a critical value, a one-Hertz signal would be detectable at -1.5 db, a

---

\* The derivation of the distribution of the variance ratio for exponentially distributed random variables is being investigated (Sogliero, 1980).

two-Hertz signal at -3 db, while the four-Hertz bandwidth signal would be detected at -6 db. It is understood that the signal is present the full 20 seconds.

Figure A3.9.2(b) shows a plot of the row variance ratio for a broadband pulsed signal. The signal has a bandwidth of 40 Hz and lasts for 1, 2, or 4 seconds. Here the effect of signal width is even more striking than in Figure A3.9.2(a). If a variance ratio of 1.5 is considered as the threshold, a 1-second pulse would be detectable at about -2 db, a 2-second pulse at about -4 db, and a 4-second pulse at about -5.5 db.

Figure A3.9.2(c) shows the plot of the column variance ratio for a slowly drifting narrowband signal. While the general trend remains as in A3.9.2(a and b) the increase in detectability is not as great due to the splitting of the variation in the data between the column and interaction sum of squares. Nevertheless, the improvement with signal width is sizeable.

Figure A3.9.2(d) shows the plot of the interaction variance ratio, the term most logically suited for the drifting narrowband signal. Unfortunately, ANOVA fails to pick up the 1-Hz signal at any SNR. If the matrix were square the results would be better. For signals of 2-Hz and 4-Hz bandwidth the results are better. However, for this configuration of matrix size and type of signal, ANOVA does not do well.

Figures A3.9.3(a) to A3.9.3(d) show the effect of frequency averaging upon the variance ratio. In Figure A3.9.3(a), a non-drifting narrowband signal shows consistent improvement in its detectability as a result of increased frequency averaging. The improvement occurs as a result of the smaller residual variance brought about by the averaging. In general,

the residual variance is reduced by a factor of  $1/N$ , where  $N$  is the number of frequencies averaged. For larger time x frequency matrices, an even greater improvement in detection might be expected.

A broadband pulsed signal displays a more complicated response as reflected in the row variance ratios plotted in Figure A3.9.3(b). While the 8 Hz averaging is highly beneficial for  $\text{SNR} > 0$  db, a lessening of its effectiveness is detected at  $\text{SNR} < 0$ . In fact at -12 db, although mostly noise effects dominate, the 8 Hz averaging fares worse than either 2 or 4 Hz averaging.

Figure A3.9.3(c) depicts the effect of frequency averaging upon a slowly drifting narrowband signal. Here, as in A3.9.3(a), the effect of averaging is highly beneficial, with 8 Hz averaging being consistently better than the other averaging levels.

Figure A3.9.3(d) shows how the drifting narrowband signal is improved by frequency averaging.

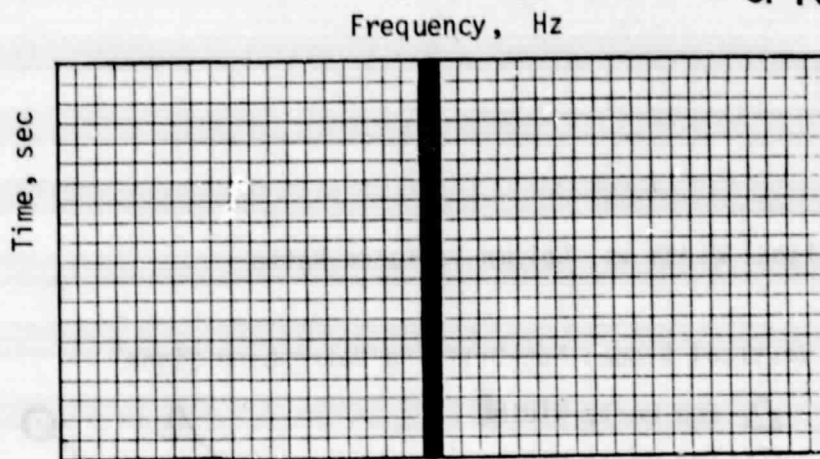
As was shown in Figure A3.9.2(d), the rapidly drifting signal is most difficult to detect for the matrix size specified--a square matrix being optimal.

Figure A3.9.4(a-d) shows the effect of bit quantization in the generation of the noise field. Quantization is a necessary outgrowth of the discretization of the data. The level of quantization refers to the number of bits that are available for the algorithm processing. For the simulation studies 2, 4, 8 and 10 bits are used.

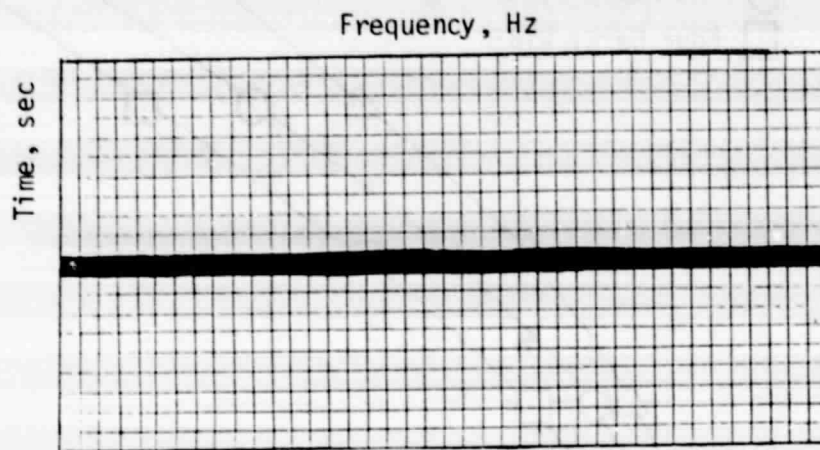
The use of 2 bits implies that the observations of the noise field are represented by 4 numbers only. This represents the number of intervals allowed in the partitioning of the exponential distribution.

It is apparent that some information is lost at this level of quantization. In (Chen, 1980) the loss of such information as a function of bit size is explored. Different scale parameters for adjusting the variance (of the exponential distribution) to the word length in order to minimize the loss of information, are presented. These values are used to adjust the quantization levels of the clipped exponential distribution.

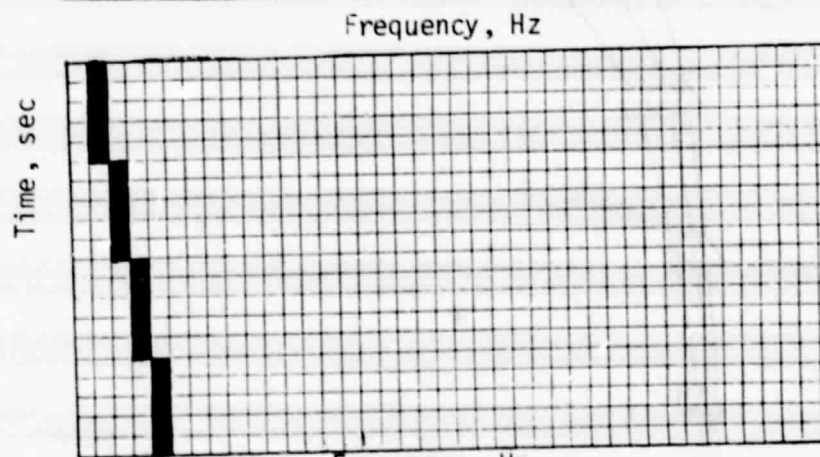
Figure A3.9.4 (a-d) reflects the results of quantizing the data of exponentially distributed noise pulse signal to bit values of 2, 4, 8 and 10. The degradation of the detectability of the signal at low quantization levels is apparent. A gradual improvement results as the bit size increases. For bit size of 10 the curves are almost identical to the non-quantized values. Figure A3.9.4 (d) shows that quantization does not improve this signal type possibility of detection. For adequate performance of the ANOVA algorithm, eight bits are recommended.



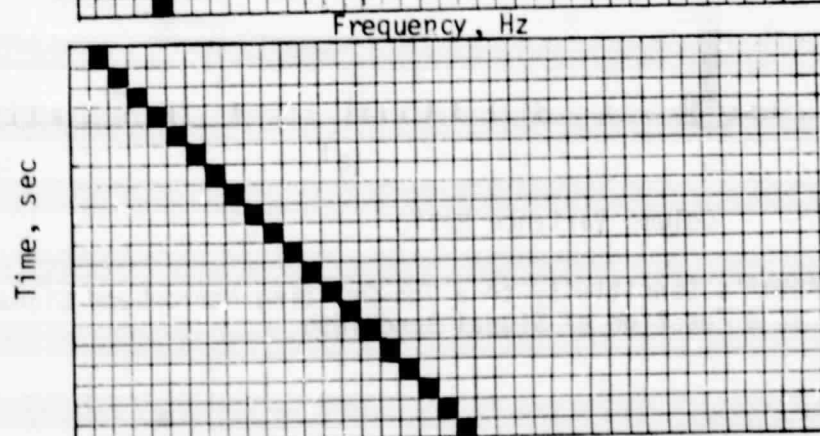
(a) Non-Drifting  
Narrowband Signal



(b) Broadband  
Pulsed Signal



(c) Slowly Drifting  
Narrowband Signal



(d) Drifting  
Narrowband Signal

Figure A3.9.1 - Signal Types Used in ANOVA Simulations



# Signal Width vs. Column Variance Ratio

Vertical Slope: Non-Drifting Narrowband Signal

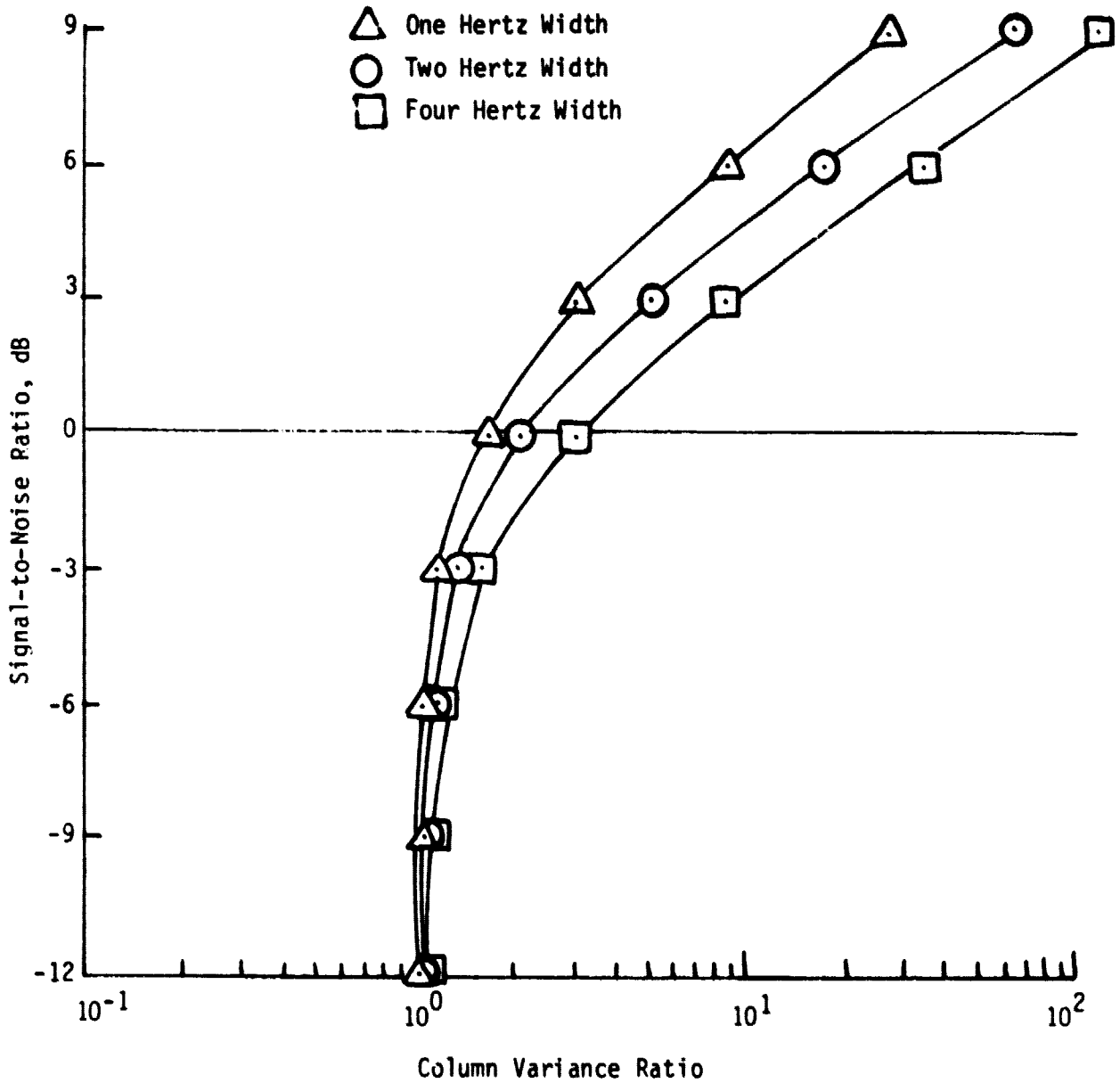


Figure A3.9.2(a) - ANOVA's Sensitivity to a Non-Drifting Narrowband Signal as a Function of Signal Bandwidth

# Signal Width vs. Row Variance Ratio

Horizontal Slope: Broadband Pulsed Signal

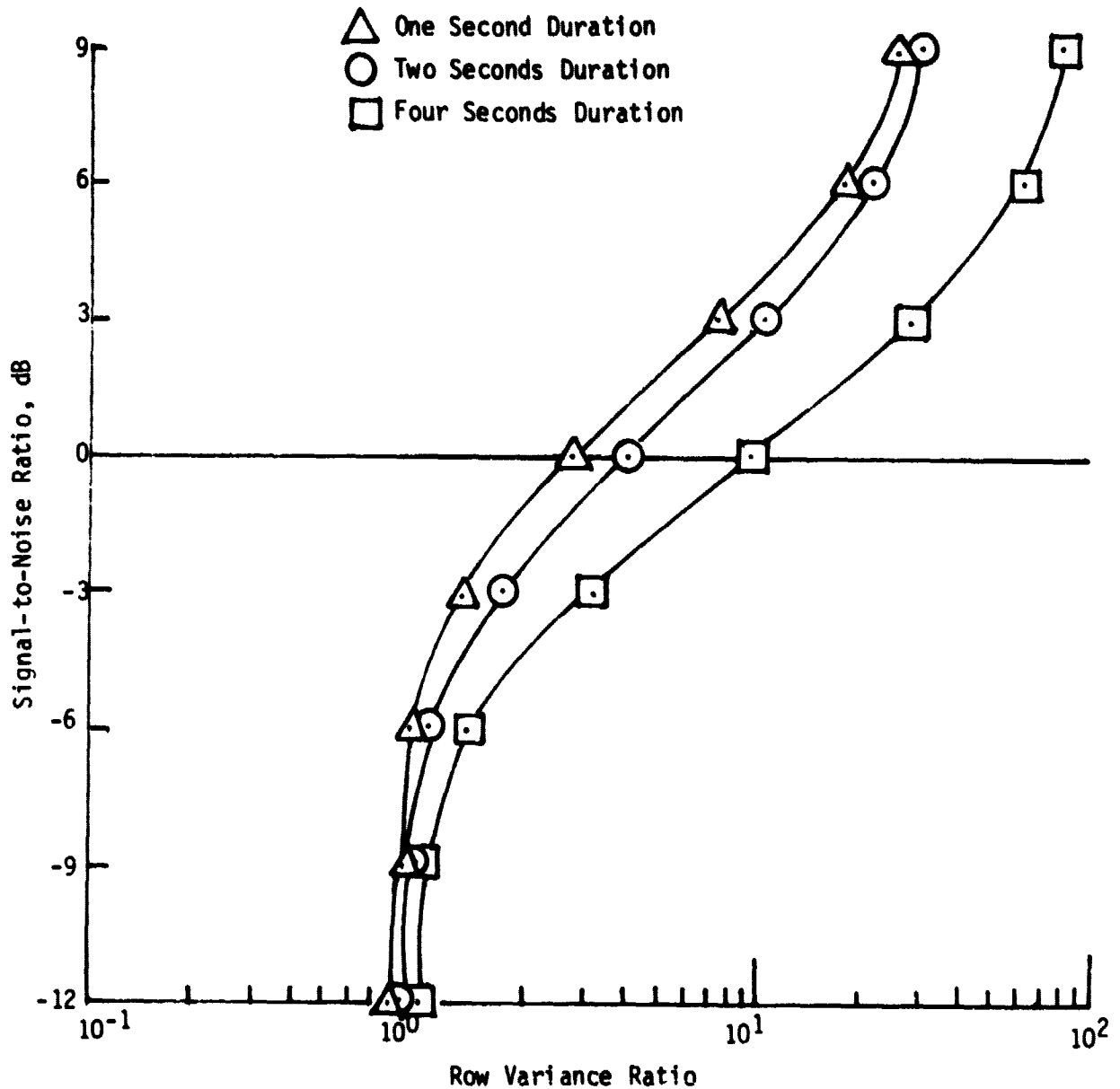


Figure A3.9.2(b) - ANOVA's Sensitivity to a Broadband Pulsed Signal as a Function of Time Duration

Slowly Drifting Narrowband Signal

△ One Hertz Width

○ Two Hertz Width

□ Four Hertz Width

Drift Rate: 0.2 Hz/sec

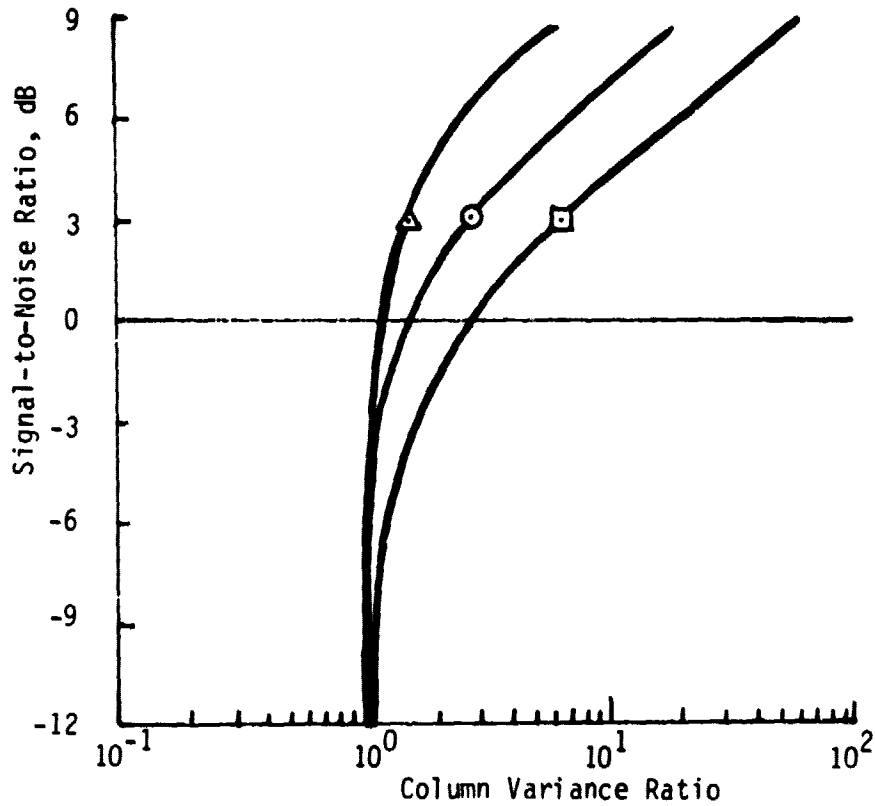


Figure A3.9.2(c) - ANOVA's Sensitivity to a Slowly Drifting Signal as a Function of Signal Bandwidth

# Drifting Narrowband Signal

- △ One Hertz Width
- Two Hertz Width
- Four Hertz Width

Drift Rate: 1 Hz/sec.

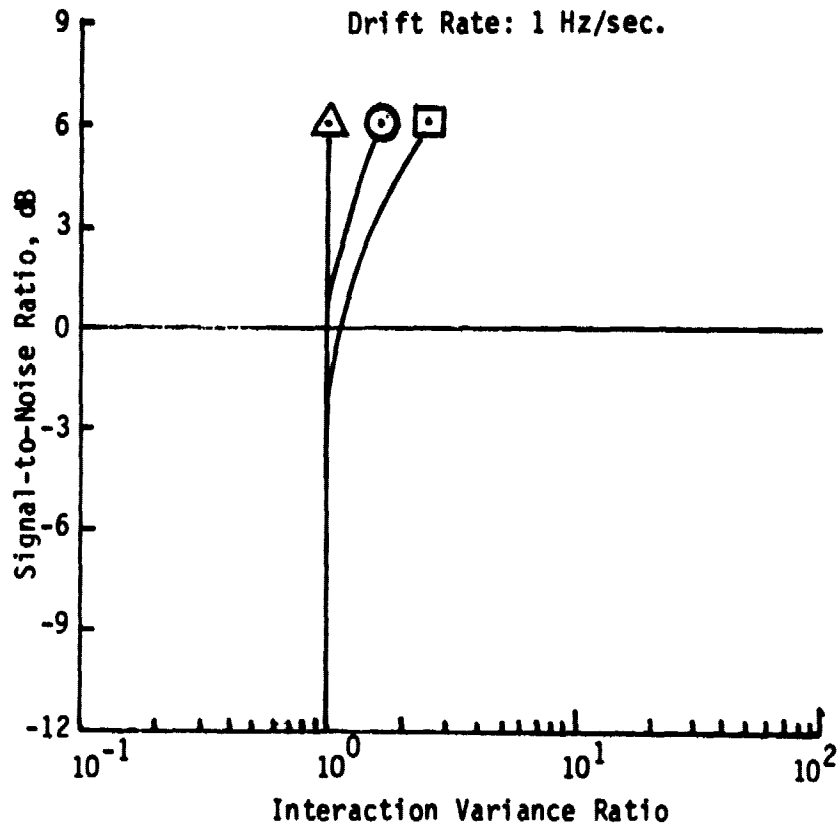


Figure A3.9.2(d) - ANOVA's Sensitivity to a Drifting Narrowband Signal as a Function of a Signal Bandwidth

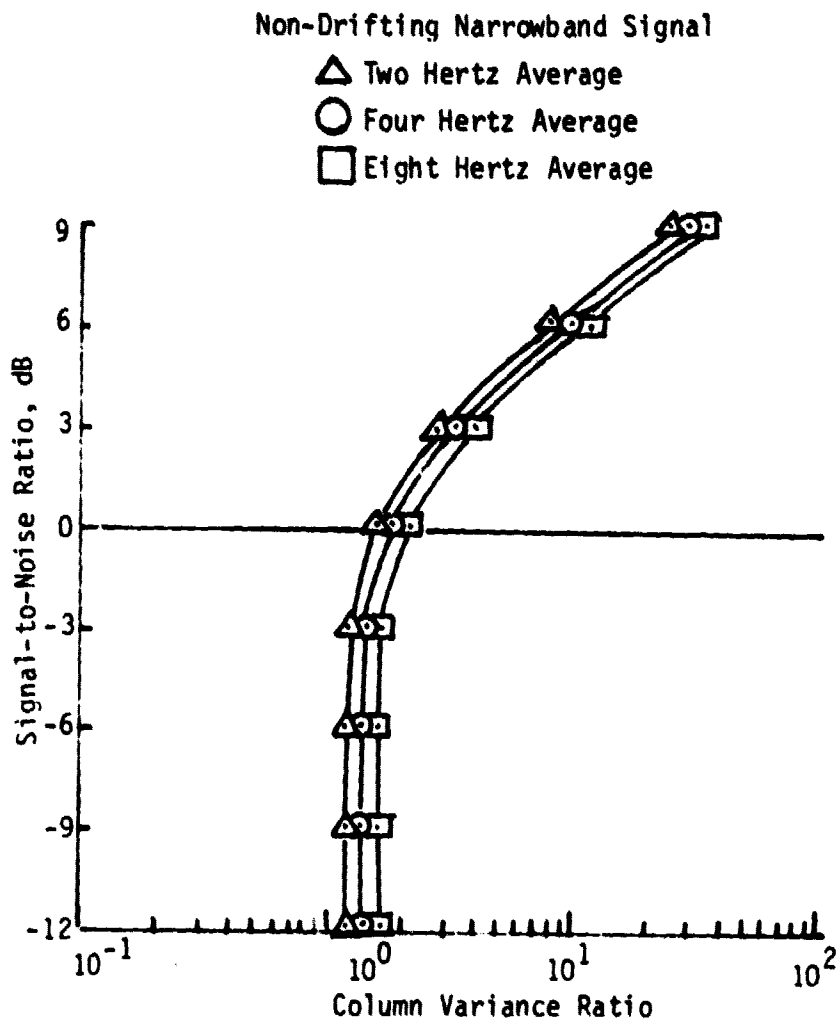


Figure A3.9.3(a) - ANOVA's Sensitivity to a Non-Drifting Narrowband Signal as a Function of Frequency Averaging

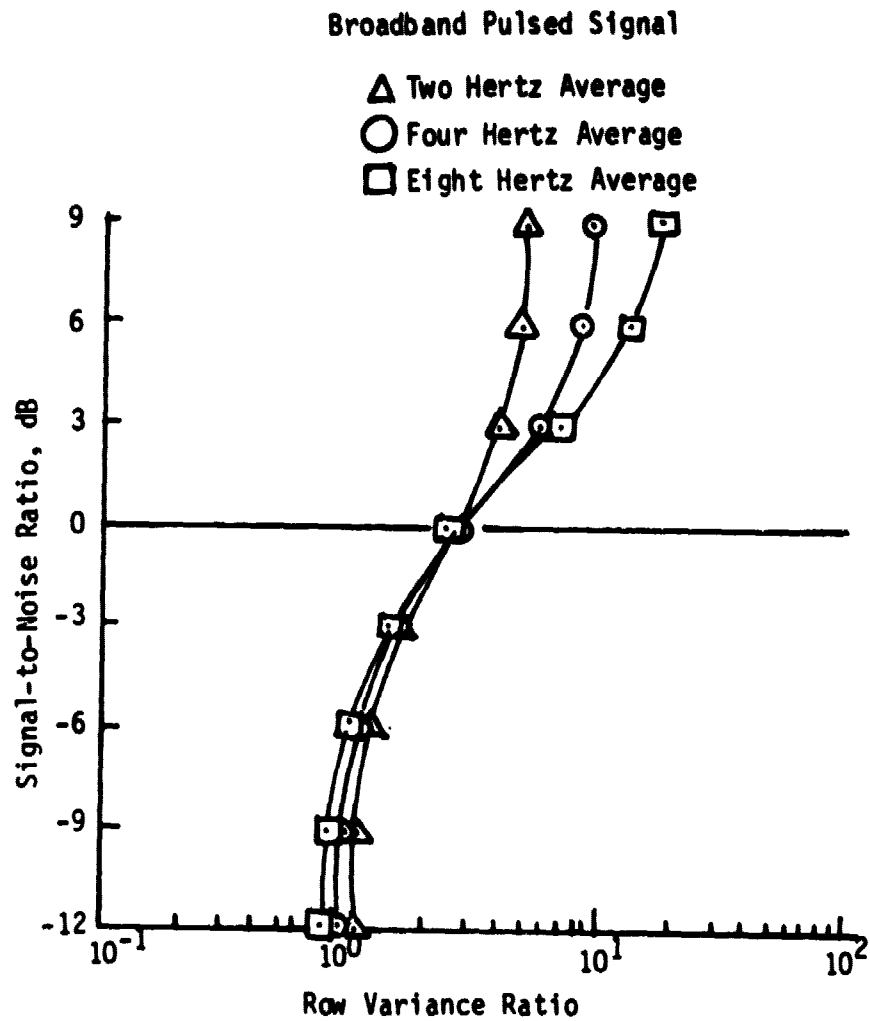


Figure A3.9.3(b) - ANOVA's Sensitivity to a Broadband Pulsed Signal as a Function of Frequency Averaging

### Slowly Drifting Narrowband Signal

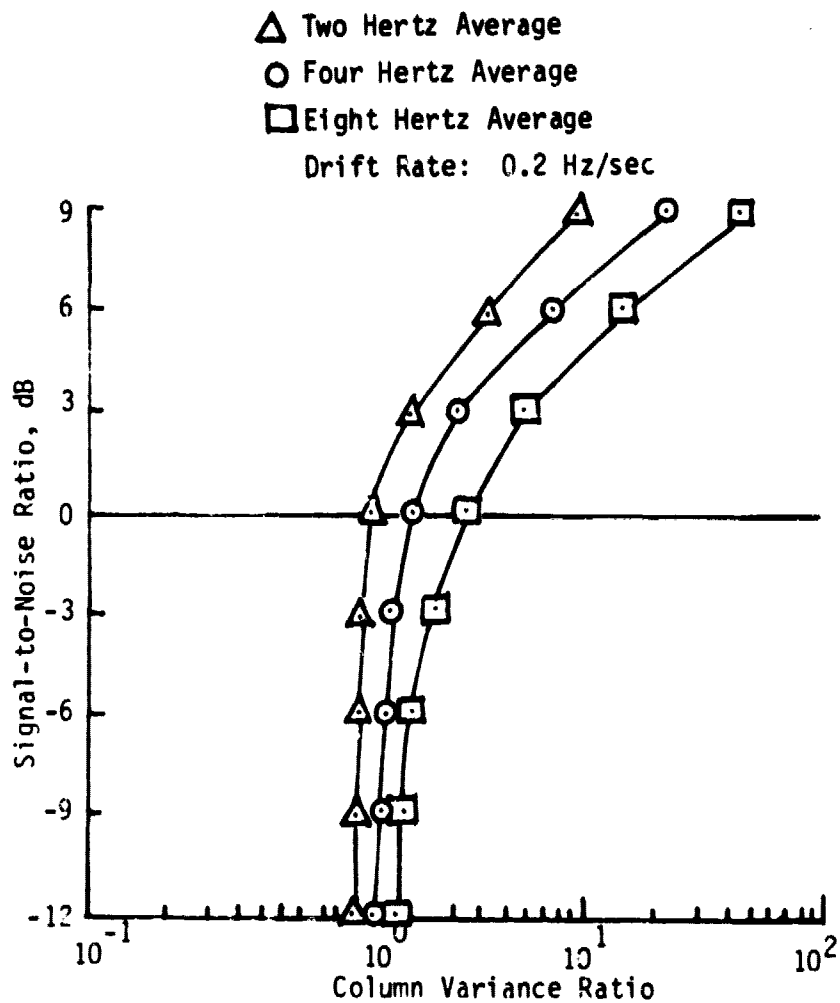


Figure A3.9.3(c) - ANOVA's Sensitivity to a Slowly Drifting Narrowband Signal as a Function of Frequency Averaging

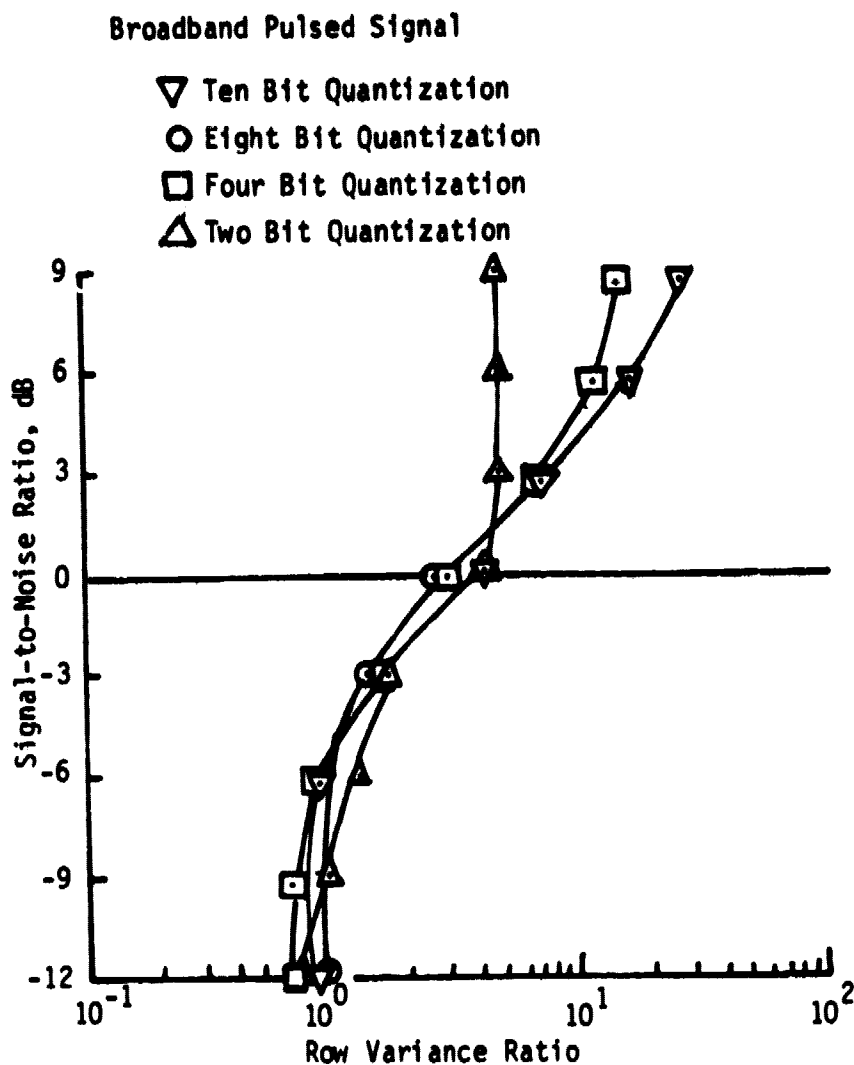


Figure A3.9.4(b)- ANOVA's Sensitivity to a Broadband Pulsed Signal as a Function of Quantization

315- 316  
 PRECEDING PAGE BLANK NOT FILMED



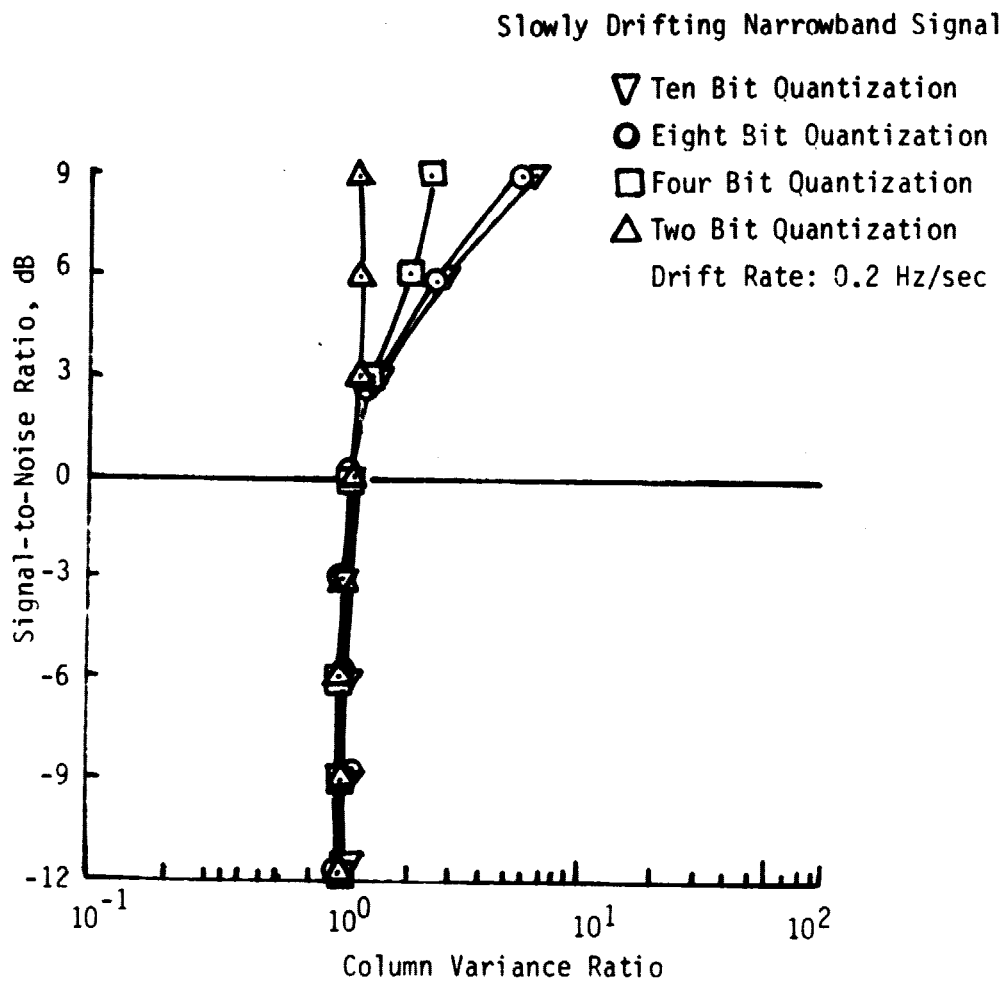


Figure A3.9.4(c) - ANOVA's Sensitivity to a Slowly Drifting Signal as a Function of Quantization

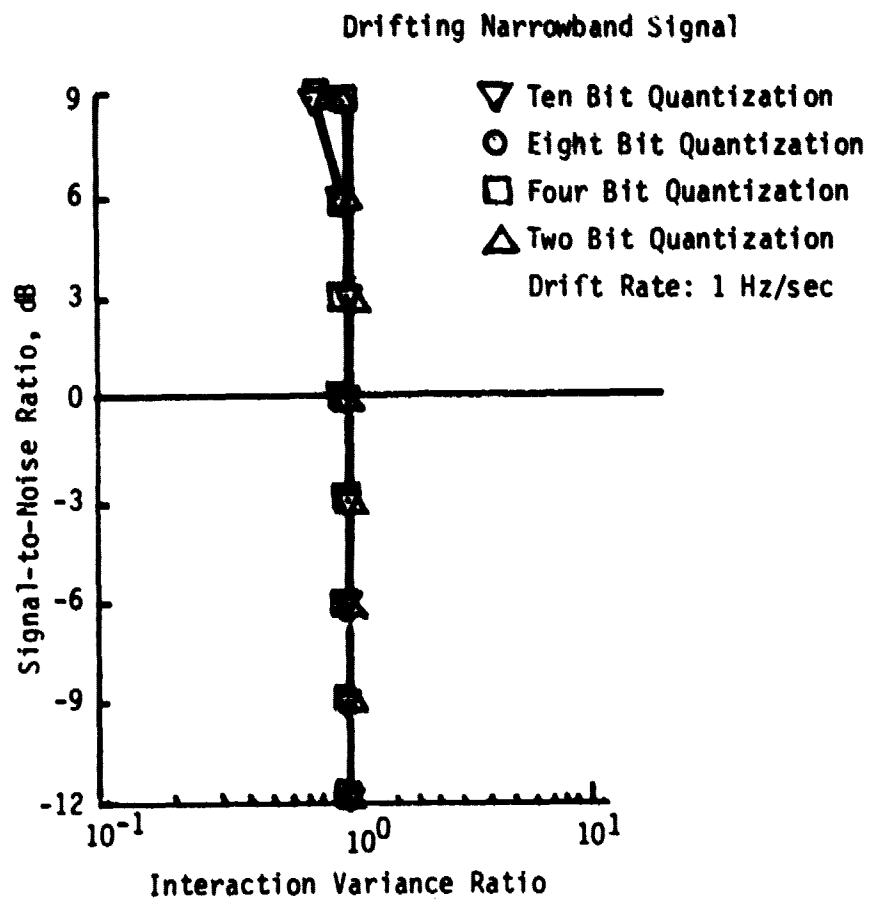


Figure A3.9.4(d) - ANOVA's Sensitivity to a Drifting Signal as a Function of Quantization

I do loathe explanations.

--Sir James Matthew Barrie,  
My Lady Nicotine

#### Appendix A3.10

#### Carrier Wave Detector Weighting Factors and Approximate Detection Statistics

##### Amplitude Weighting and Signal Distribution Models

The weighting procedure described below is a way of applying techniques of the Radon Transform (Radon, 1917) to the signal extraction problem at hand. Since it allows a signal to be picked up in a continuous, linear fashion, it is in some respects a generalized digital version of the analog, delay line video detector described in the CYCLOPS Report (1973, pp. 132-133). In general there are three potential signal distribution modes that could be used to determine weighting factors: the delta-function model, which uses the value in the frequency bin whose center falls within the ray; the uniform distribution model, which uses weighting factors proportional to the area of each frequency bin contained within the ray; and the sinc distribution model which uses weighting factors proportional to the area under the sinc curve.

The MCSA impresses a sinc weighting function on the incoming data; therefore, the delta-function model is not exactly appropriate for this situation.

The most obvious choice is to use exactly the same model as the MCSA; however, the following analysis of the difference between the two models seems to indicate that little is gained by going to a sinc model, and a uniform model is much easier to implement.

Consider a single time-frequency "picture element" (pixel) of the array with unit length and unit width. The percentage of the pixel value (pv) contributed to the total ray sum depends upon the way the pv is distributed throughout the pixel and the way in which the pixel is overlapped by the ray. For the uniform model the percentage is just that portion of area inside the ray, which in Figure A3.10.1 is just  $1 \cdot x = x$ ; the contribution to the ray sum is  $x$  times pv.

Consider the case in which the weighting is by the sinc function which is, by definition (Bracewell, 1978),

$$\text{sinc } x = \frac{\sin \pi x}{\pi x}$$

And, suppose the pv is distributed over the pixel so that the zero crossings occur at zero power, corresponding to the corners of the pixel, as illustrated in Figure A3.10.2. Since we wish to include the central lobe in the pixel, our purposes are served best by considering  $\text{sinc}(2x - 1)$ . In the sinc model the contribution to the ray sum made by a given pixel is given by  $f(x)$  times pv where  $f(x)$  is

$$f(x) = \frac{\int_0^x \text{sinc}(2t - 1) dt}{\int_0^1 \text{sinc}(2t - 1) dt} \quad (\text{A3.10-1})$$

Thus,  $f(x)$  is an appropriately normalized fraction of area under the sinc function.

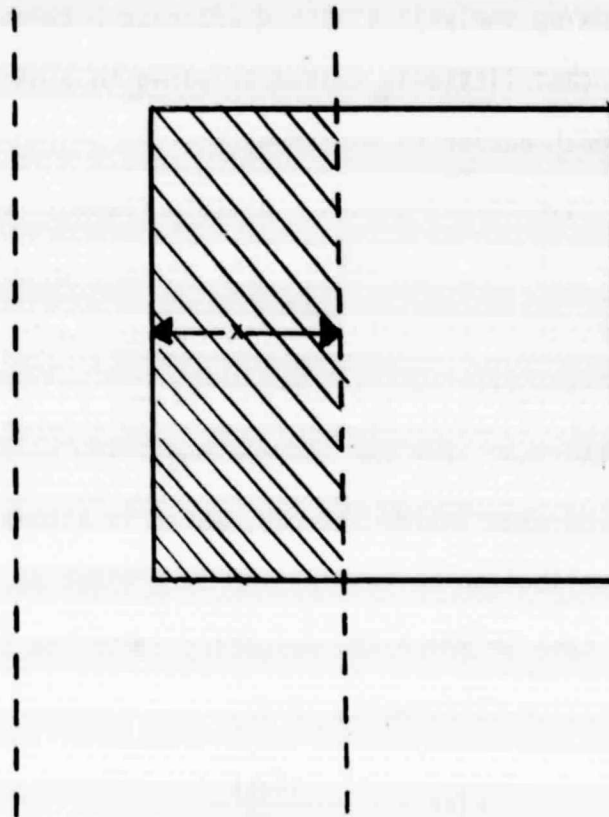


Figure A3.10.1 - In the uniform model the contribution to the ray sum is just that portion of the pixel which lies inside the ray multiplied by the pixel value. The variable  $x$  represents the fraction of pixel overlapped by the ray.

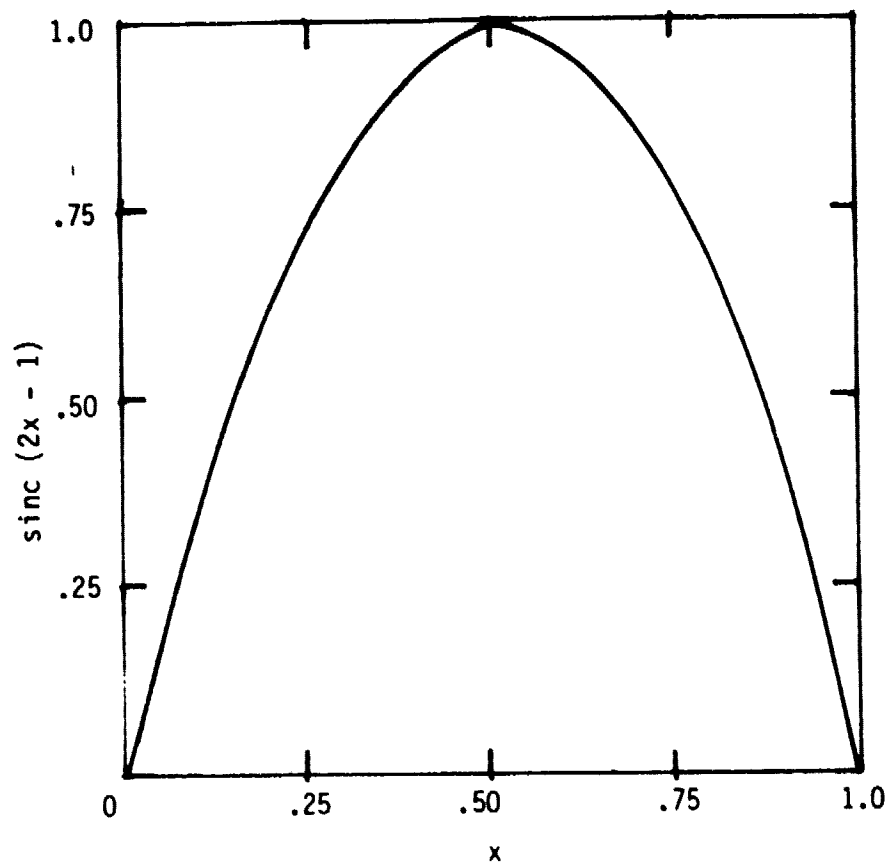


Figure A3.10.2- The Central Lobe of the Sinc Function

For purposes of calculation it will be useful to simplify (A3.10-1)  
 First, observe the defining relation for the Sine Integral function  
 (Abramowitz and Stegan, 1964).

$$\text{Si}(x) = \int_0^x \frac{\sin t}{t} dt \quad (\text{A3.10-2})$$

By (A3.10-2) it is straightforward to obtain

$$\text{Si}(\pi x) = \pi \int_0^x \text{sinc } t \, dt, \quad (\text{A3.10-3})$$

and

$$\int_0^x \text{sinc}(2t - 1) \, dt = \frac{1}{2\pi} \text{Si}(\pi) + \int_0^{2x-1} \text{sinc } t \, dt \quad (\text{A3.10-4})$$

For  $0 < x < 1/2$  the integral on the right side of (A3.10-4) is negative  
 it is useful to take that into consideration explicitly; thus

$$\begin{aligned} \text{sinc}(2t - 1) \, dt &= \frac{1}{2\pi} \text{Si}(\pi) - \frac{1}{2\pi} \text{Si}(\pi - 2\pi x), \quad 0 \leq x \leq 1/2 \\ &\frac{1}{2\pi} \text{Si}(\pi) + \frac{1}{2\pi} \text{Si}(2\pi x - \pi), \quad 1/2 \leq x \leq 1 \end{aligned} \quad (\text{A3.10-5})$$

We now observe that the denominator in (A3.10-1) is  $\frac{1}{\pi} \text{Si}(\pi)$  and rewrite  
 $f(x)$  in terms of the tabulated Sine Integral function,

$$f(x) = \begin{cases} \frac{1}{2} - \frac{1}{2\text{Si}(\pi)} \text{Si}(\pi - 2\pi x), & 0 \leq x \leq 1/2 \\ \frac{1}{2} + \frac{1}{2\text{Si}(\pi)} \text{Si}(2\pi x - \pi), & 1/2 \leq x \leq 1, \end{cases} \quad (\text{A3.10-6})$$

$$\text{where } \text{Si}(\pi) = 1.8519370. \quad (\text{A3.10-7})$$

The values for  $f(x)$  were computed for  $0 \leq x \leq 1$  and were used to  
 compute the per cent of the pv included in the ray sum which is given by  
 $f(x)$  times 100. In Figure A3.10.3 the results are plotted as a function of the  
 fraction of overlap  $x$ . The straight line represents the percent of the

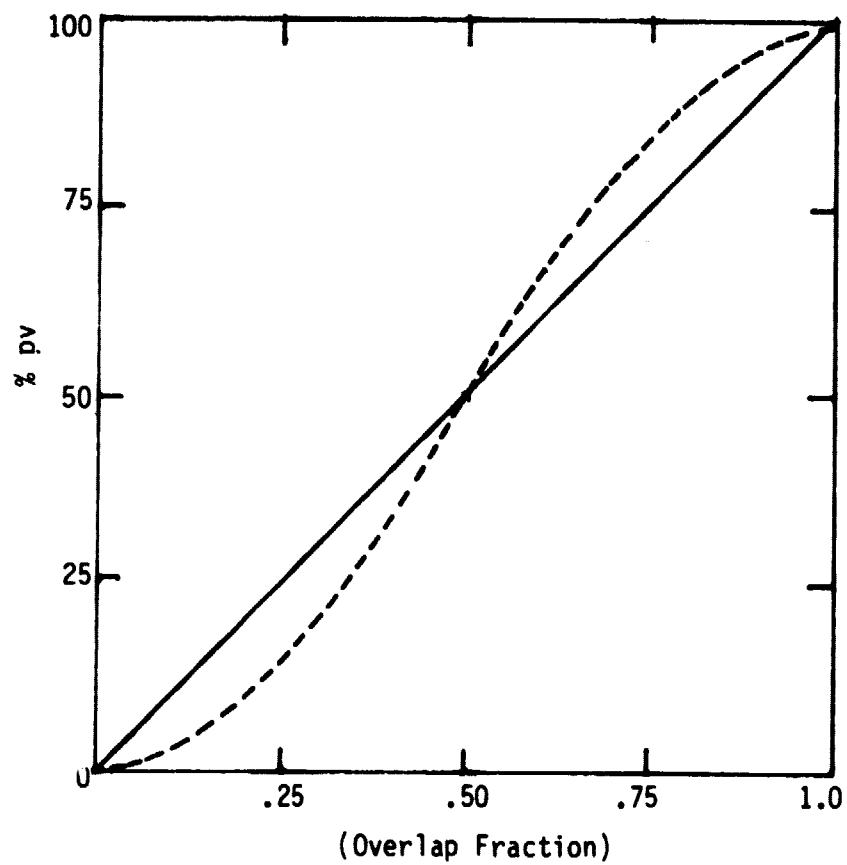


Figure A3.10.3 - Comparison of Uniform (solid line) and Sinc (dashed line) Models



pixel value included in the ray sum for the uniform model and is included for comparison with the graph for the sinc model.

A moment of thought reveals that the results of Figure A3.10.3 hold even if the ray is not parallel to the pixel, provided  $x$  is normalized to represent the fraction of overlap for the more general case.

In the case where the pv is distributed over the pixel as indicated in Figure A3.10.4 so that the zero crossings occur at half-power, corresponding to the midpoints of adjacent pixels, the function  $f(x)$  is modified to

$$f(x) = \frac{\int_0^x \text{sinc}(t - 1/2) dt}{\int_0^1 \text{sinc}(t - 1/2) dt} \quad (\text{A3.10-8})$$

In terms of the Sine Integral function,

$$f(x) = \begin{cases} \frac{1}{2} - \frac{1}{2\text{Si}(\pi/2)} \text{Si}(\pi/2 - \pi x), & 0 \leq x \leq 1/2 \\ \frac{1}{2} + \frac{1}{2\text{Si}(\pi/2)} \text{Si}(\pi x - \pi/2), & 1/2 \leq x \leq 1 \end{cases} \quad (\text{A3.10-9})$$

where

$$\text{Si}(\pi/2) = 1.3707622 \quad (\text{A3.10-10})$$

This "second" sinc model is shown in Figure A3.10.5 (dashed line) where it can be compared with the uniform model and the previous sinc model. Observe that for the second sinc model there is very little deviation from the uniform.

The weighting function used by the MCSA is most closely matched by the second sinc model and therefore, seems to be the most logical choice. However, since the difference between it and the uniform model is so slight it is recommended from the standpoint of the implementation that the weighting factors for the ray be based on the uniform model.

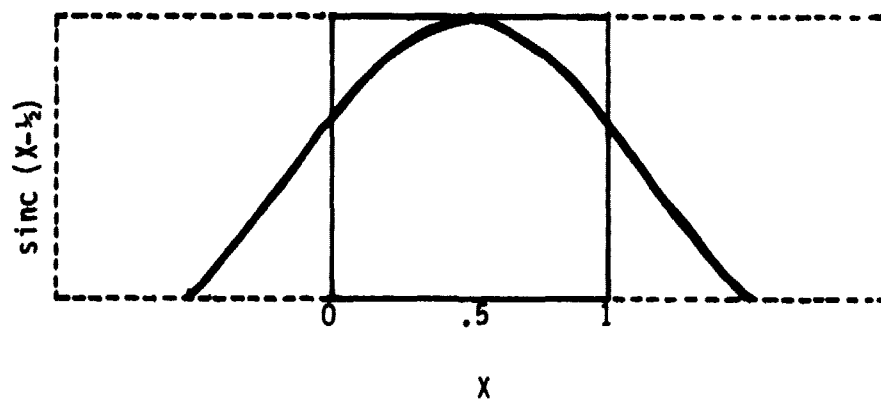


Figure A3.10.4 - Sinc ( $X - \frac{1}{2}$ ) as a Function of  $X$

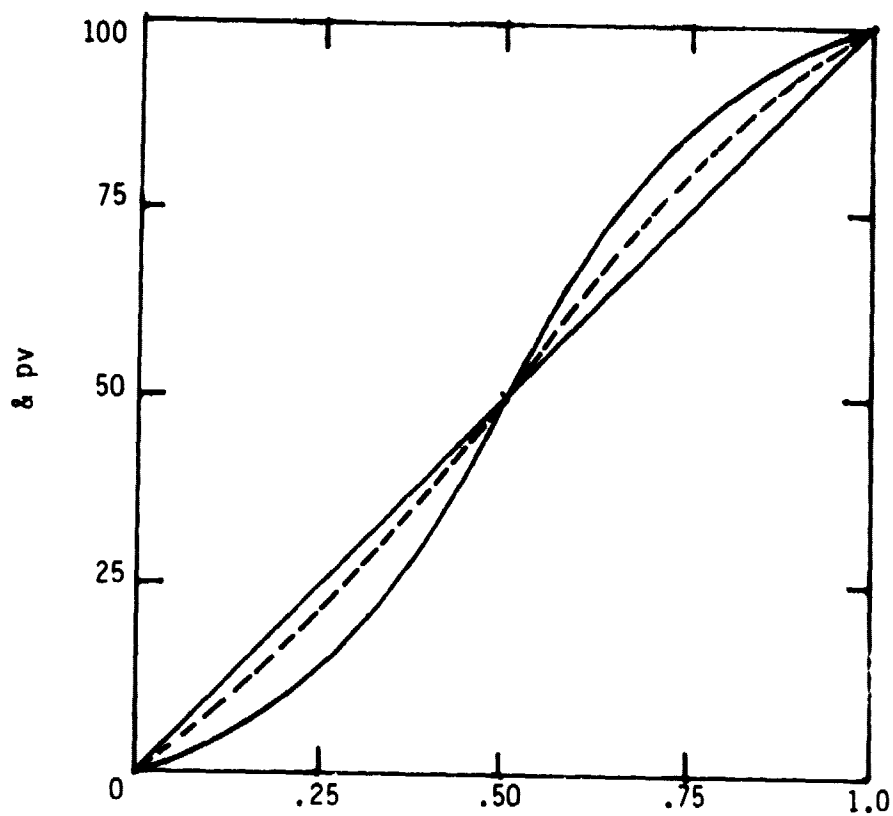


Figure A3.10.5 - Comparison Among Uniform (straight solid line), Second Sinc Model (dashed line), and First Sinc Model (curved solid line)

The special way in which the Radon transform (Radon, 1917; Gel'fard, 1966) is applied to the GCV numbers or power spectrum for signal enhancement and linear pattern selection has been shown (Deans, 1979) to be equivalent to a transform known as the Hough transform to people in the area of pattern analysis and image processing. This transform grew out of early attempts to implement automatic and rapid methods for analyzing particle tracks in bubble chamber photographs (Hough and Powell, 1961; Hough, 1962; Franck, Hough, and Powell, 1963). The basic idea is to designate the frequency-time space by the rectangle xy coordinate system, see Figure A3.10.6. Collinear points in this space may be characterized by the line

$$\rho = x \cos \theta + y \sin \theta$$

This approach is used to map collinear points in the xy plane to single cell regions in a space with rectangular coordinates ( $\theta, \rho$ ) which we call parameter space (see Figure A3.10.6(b)).

#### Calculation of Ray Amplitude Factors

Usually, the discrete implementation of Radon transform is performed using line integrals or ray sums, where the rays are equal in width to the width of the pixel. Implementation constraints here dictate that the ray widths vary as a function of angle. (The horizontal distance is kept constant and equal to one unit as illustrated in Figure 3.9.) A method is presented for computing the ray factors which does not require that the ray width be the same as the pixel width. A sample calculation is given to illustrate the method for finding the ray factors associated with a given cell or pixel.

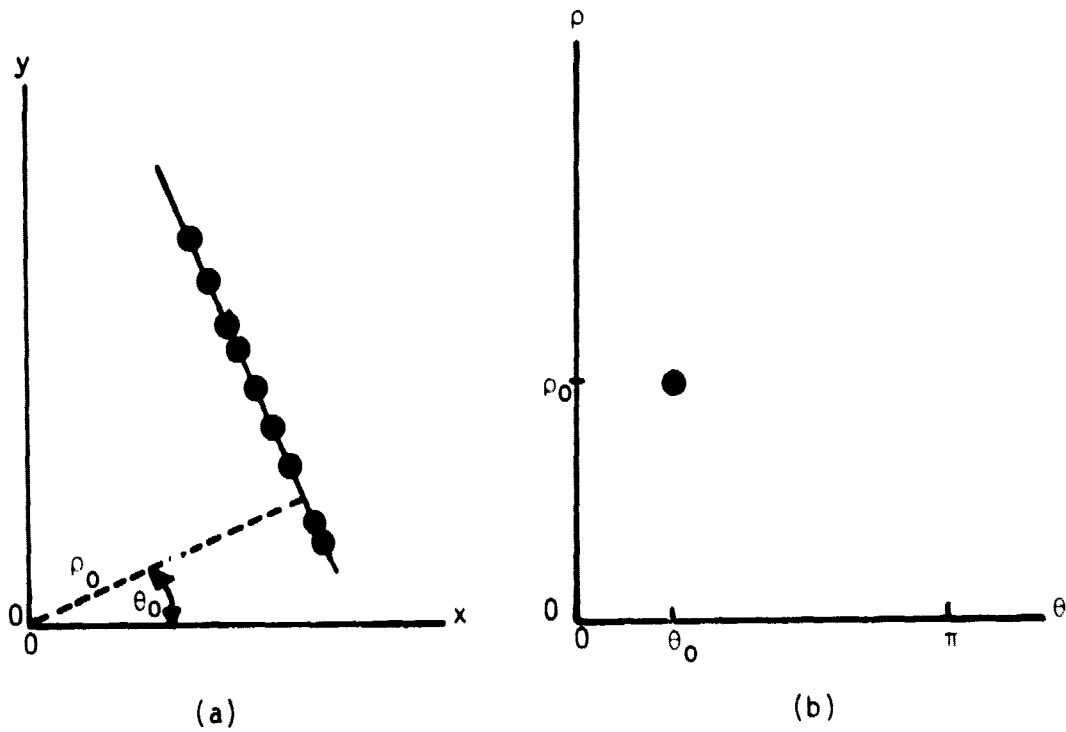


Figure A3.10.6 - Map from Collinear Points in the  $xy$  Plane to a Single Cell in Parameter Space. The Hough Transform.

Given a unit cell (or pixel) at location  $(i+1,j+1)$  in the grid indicated in Figure A3.10.7 it is desired to find the line factors and ray factors associated with the cell. The line factors may be found directly from the Radon transform of the characteristic function

$$F(x,y) = \chi(i+1,j+1) = \begin{cases} 1 & \text{whenever } (x,y) \in C_{i+1,j+1} \\ 0 & \text{elsewhere} \end{cases}$$

In Figure A3.10.7 the line factor is just the length of the line segment AB (shown in inset). The ray factor is the area of the shaded region. Once the line factors are determined it is easy to calculate the ray factors by adding areas, see Figure A3.10.8.

The line factors for  $C_{i+1,j+1}$  are found from the Radon transform formula

$$f_j = \iint_D \chi(i+1,j+1) \delta(\theta_k - x\cos\theta - y\sin\theta) dx dy$$

where the integral over the domain  $D$  reduces to an integral over  $C_{i+1,j+1}$ .

We now give a prescription for finding the various factors. Details of derivation may be found elsewhere (Deans, 1979).

- (1) Given  $\theta$  and  $\rho_k$ :

Calculate  $\rho = \rho_{ij} = \rho_k - i\cos\theta - j\sin\theta$

If  $\rho$  is not in the region  $0 < \rho \leq \sin\theta + \cos\theta$  the line does not pass through the cell and the line factor is zero.

- (2) Use Table A3.10.1 to calculate the line factors,  $f_k$ .
- (3) Use Table A3.10.1 to calculate the ray factors,  $r_k$ .
- (4) Once the ray factors are calculated there are various cases to consider in order to find the contribution made by the cell under consideration to the total ray sum for various rays.

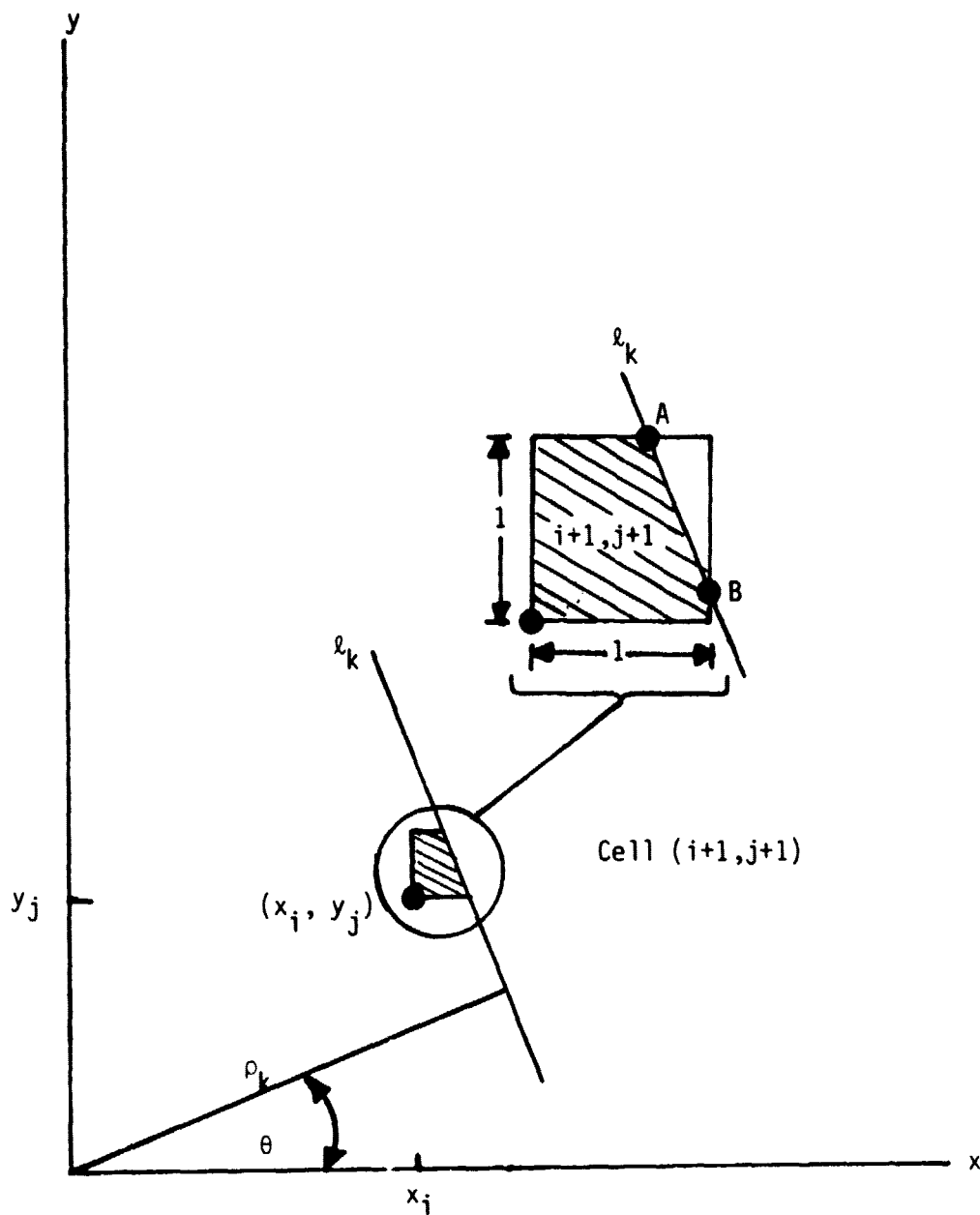


Figure A3.10.7 - A Unit Cell. The Length of Line AB through the Cell is the Line Factor, and the Shaded Region is the Ray Factor.

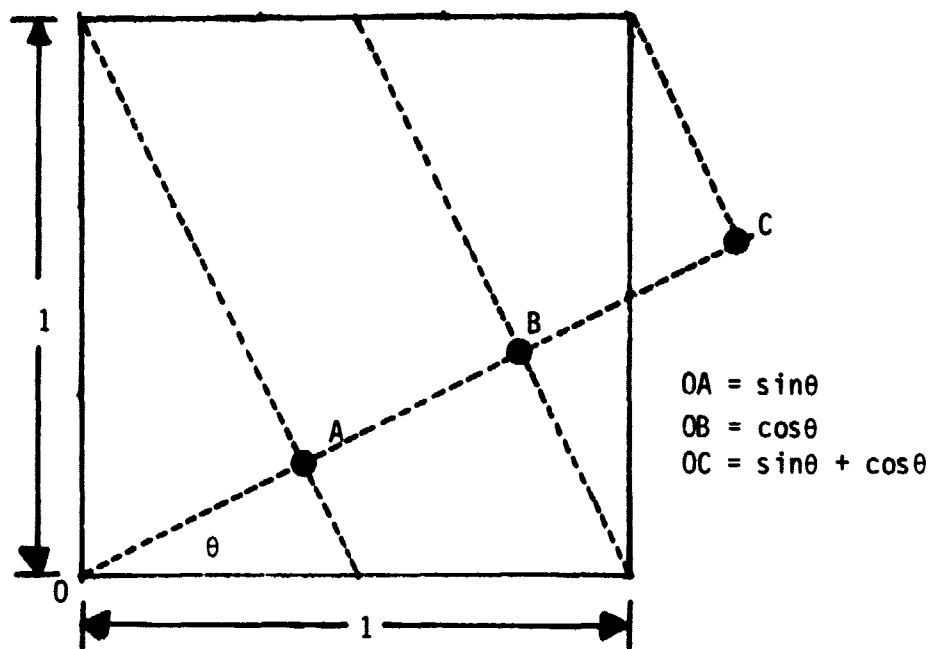


Figure A3.10.8 - Geometry for Ray Factors



Region for $\rho$	line factor $f_k$	ray factor $r_k$
$0 < \rho \leq \sin \theta$	$\frac{\rho}{\sin \theta \cos \theta}$	$\frac{\rho^2}{\sin 2\theta}$
$\sin \theta < \rho < \cos \theta$	$\frac{1}{\cos \theta}$	$\frac{2\rho \sin \theta - \sin^2 \theta}{\sin 2\theta}$
$\cos \theta \leq \rho \leq \sin \theta + \cos \theta$	$\frac{\sin \theta + \cos \theta - \rho}{\sin \theta \cos \theta}$	$\frac{2\rho(\sin \theta + \cos \theta) - (1 + \rho^2)}{\sin 2\theta}$

Table A3.10.1 - Line and Ray Factors

For the ray width equal to unity (same as cell width) there are only two cases (a) and (b), illustrated in Figure A3.10.9.

- (a) Cell  $i, j$  contributes to ray sum  $S_k$  and  $S_{k+1}$ . The contribution to  $S_k$  is just the ray factor  $r_k$  for the line  $\ell_k$ . The contribution to  $S_{k+1}$  is one minus the ray factor,  $1 - r_k$ .
- (b) In this case a contribution is made to  $S_k$ ,  $S_{k+1}$  and  $S_{k+2}$ . The contribution to  $S_k$  is  $r_k$ . The contribution to  $S_{k+1}$  is  $r_{k+1} - r_k$ . The contribution to  $S_{k+2}$  is  $1 - r_{k+1}$ .

A sample calculation is shown for the indicated cell in Figure A3.10.10.

Calculation for C(21,7):

Step 1: C(21,7) contributes to rays 22 and 23.  $\theta = 14.04^\circ$

$$\rho_k = \rho_{22} = 22 \cos \theta$$

$$\rho = \rho_k - 20 \cos \theta - 6 \sin \theta = 21.343 - 20.858$$

$$\rho = 0.485$$

Step 2: Contribution to ray 22:

$$r_{22} = \frac{2\rho \sin \theta - \sin^2 \theta}{\sin^2 \theta} = .375$$

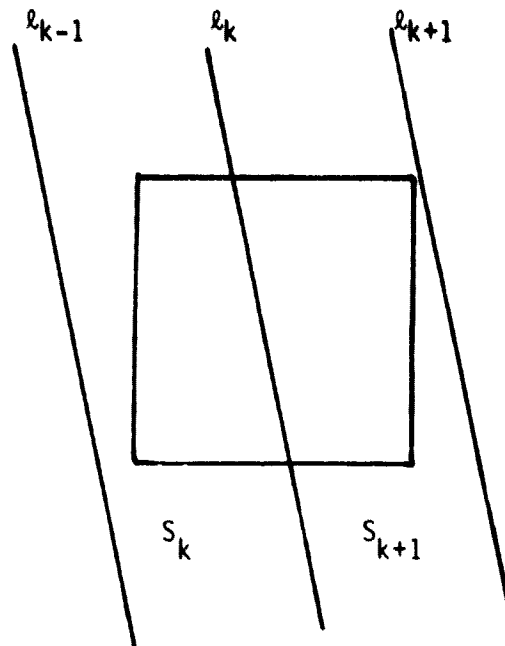
Contribution to ray 23:

$$r_{23} = 1 - r_{22} = .625$$

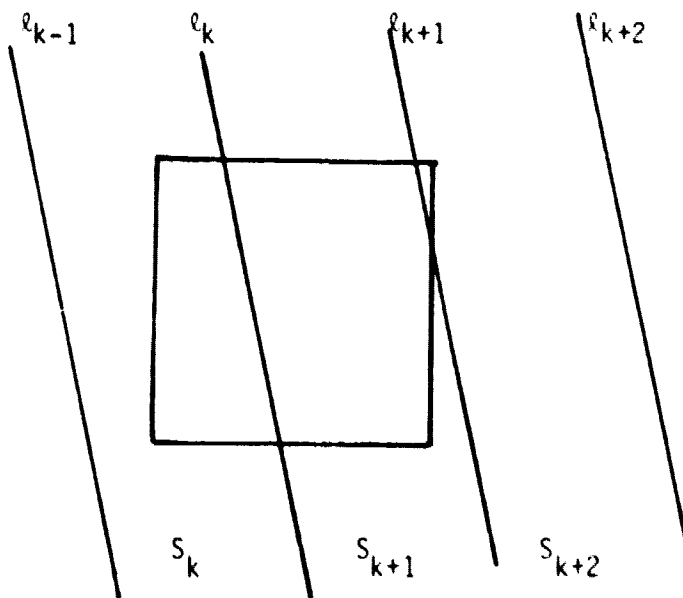
## Phase Factors

### Overview

There are two separate phase corrections that must be applied in order to properly combine adjacent time-frequency cells into the generalized coherence value along a given line. One of them is due to the phase of the signal itself, and the other is caused by phase shifts that occur in the



Case (a)



Case (b)

Figure A3.10.9 - Contributions made by Ray Factors to Ray Sums

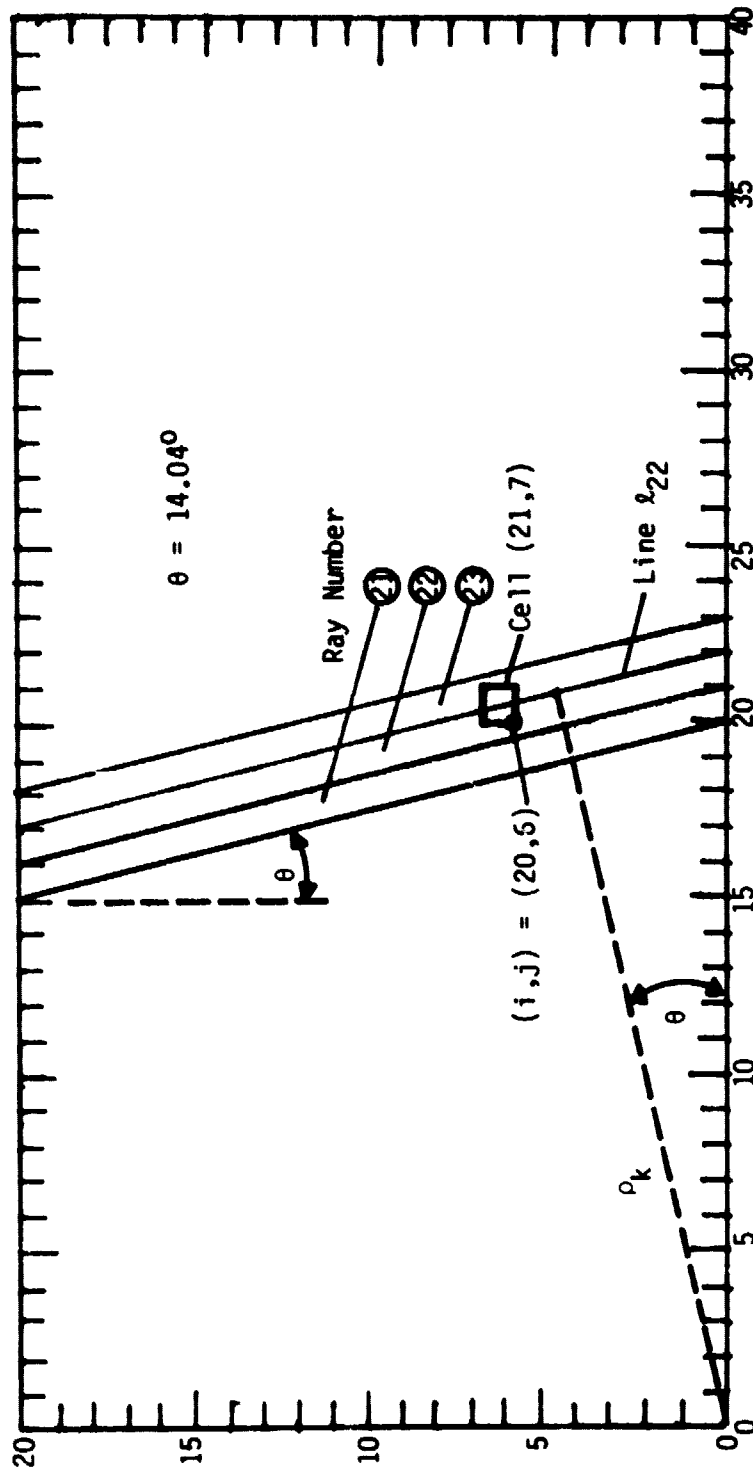


Figure A3.10.10 - Sample Calculation of Ray Factors for Cell (21, 7)

MCSA. The two corrections are independent, and can be made separately.

### Signal-Caused Corrections

From equation (A.28) we can write the polarization-matched signal as

$$z_i = A_i e^{j(\pi \dot{f} i^2 + 2\pi f_0 i + \phi_i)}$$

where

$\dot{f}$  = frequency drift rate (Hz/sec)

$f_0$  = initial frequency (when  $i = 0$ )

$\phi_i$  = initial (arbitrary) phase angle

$i$  = integer time index;  $i = 0$  initially;  $i$  increases in steps of one second.

Note that the phase is a quadratic function of time.

The generalized coherence factor involving  $z_i$  is

$$\text{GCV} \sim \left| \sum_{i=0}^{N-1} z_{i+1} z_i^* \right|$$

Using

$$z_{i+1} = A_{i+1} e^{j(\pi \dot{f}(i+1)^2 + 2\pi f_0(i+1) + \phi_{i+1})}$$

we obtain

$$\begin{aligned} \text{GCV} &\sim \left| \sum_{i=0}^{N-1} A_i A_{i+1} e^{j(\pi \dot{f}(i+1)^2 + 2\pi f_0(i+1) + \phi_{i+1}) - (\pi \dot{f} i^2 + 2\pi f_0 i + \phi_i)} \right| \\ \text{GCV} &\sim \left| \sum_{i=0}^{N-1} A_i A_{i+1} e^{j i(2\pi \dot{f})} e^{j(\pi \dot{f} + 2\pi f_0)} e^{j(\phi_{i+1} - \phi_i)} \right| \\ \text{GCV} &\sim \left| e^{j(\pi \dot{f} + 2\pi f_0)} \sum_{i=0}^{N-1} A_i A_{i+1} e^{j 2\pi \dot{f} i} e^{j(\phi_{i+1} - \phi_i)} \right| \end{aligned}$$

Since the magnitude of the phase term outside the summation is one, we have

$$\text{GCV} \sim \left| \sum_{i=0}^{N-1} A_i A_{i+1} e^{j 2\pi \dot{f} i} e^{j(\phi_{i+1} - \phi_i)} \right|$$

or

$$\text{GCV} \sim \left| \sum_{i=0}^{N-1} z_{i+1} z_i^* e^{+j2\pi f_i} \right|$$

This is the value that would be measured if no phase correction were made.

In order to cancel the phase variation, we actually compute

$$\text{GCV} \sim \left| \sum_{i=0}^{N-1} z_{i+1} z_i^* e^{-j2\pi f_i} \right| \quad (\text{A3.10-11})$$

The linear phase invariance property of generalized coherence thus requires that only a linear phase correction be made to a signal with a quadratic phase variation.

#### MCSA-Caused Phase Corrections

The phase shift introduced by the MCSA is derived in Appendix A. For a drifting signal, the phase shift depends on the path of the signal through the frequency bin during the integration period. A good approximation to the phase shift is  $\pi \Delta f_{\frac{1}{2}}$  radians, where  $\Delta f_{\frac{1}{2}}$  is the difference frequency between the signal and the center of the bin (in units of binwidths), measured halfway through the integration period. Table 3.5 compares the approximate and exact phase shift, for linear drift rates up to 1/2 Hz/sec, which is the maximum being considered for this system.

The approximate method will be used for the remainder of this discussion, because it is easier to visualize. The exact method assumes that the signal has infinitesimally small bandwidth, whereas the approximate method is somewhat more general. This comparison is analogous to the case of the amplitude weighting factors, where sinc weighting is exactly correct for infinitesimally narrowband signals, but uniform weighting

gives nearly the same result, and is applicable to wider bandwidth signals as well. Since the phase factors only need be computed once and then hardwired into the system, the exact method of Appendix A could still be used if desired, following the general plan to be described here.

<u>Drift Range</u>	<u>Exact method phase shift</u>	<u>Approximate method phase shift</u>
$0 \leq d \leq 0.5$	$41^{\circ}.4$	$45^{\circ}$
$0 \leq d \leq 0.2$	$16^{\circ}.9$	$18^{\circ}$
$0 \leq d \leq 0.1$	$8^{\circ}.1$	$9^{\circ}$

Table A3.10.2 - Comparison of Exact and Approximate MCSA Phase Shift Computations (See Appendix A)

Let us first consider phase changes which occur as a function of time (i.e., from one integration interval to the next). Changes which occur as a function of frequency between adjacent frequency bins in the same integration interval will be considered subsequently. In Figure A3.10.11, the solid lines are the edges of the Radon band, the dashed line is its center, and the dotted lines mark the halfway time point of two successive integration periods. The boxes represent time-frequency bins. Let us first assume that the signal is drifting exactly along the centerline of the radon band, at a drift rate of  $\dot{f}$ . Using the approximate phase shift method discussed above, we see that the MCSA will introduce phase shifts of  $\pi(C-B)$  in bin 1 and  $\pi(F-E)$  in bin 2. Since  $F=C+\dot{f}$  and  $E=B$ , the shift in bin 2 is also  $\pi(C-B+\dot{f})$ . Thus the shift in bin 2 differs from that of bin 1 only by  $\pi\dot{f}$ . It may also be seen that the same difference would occur between bin 3 and bin 2, and between any two successive bins in a

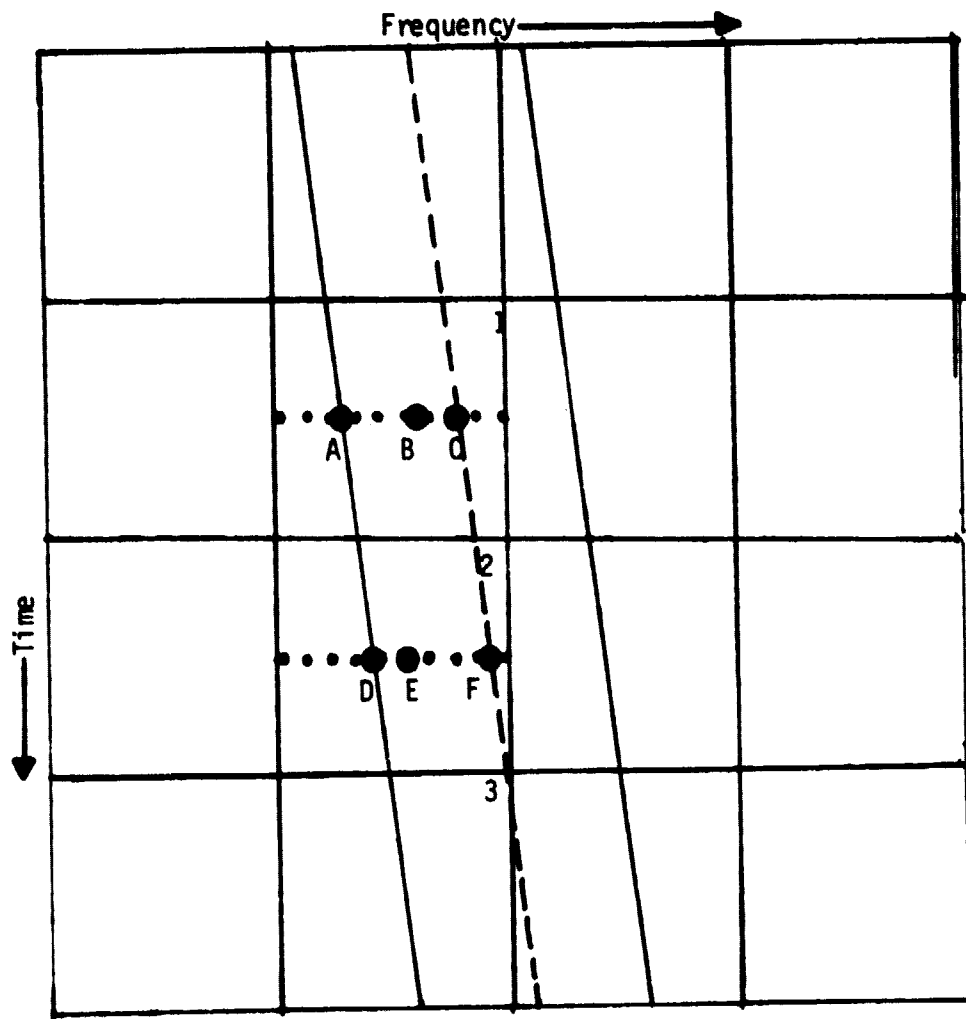


Figure A3.10.11 - MCSA Phase Shifts Between Successive Integration Intervals



vertical row. Suppose now that the signal was not at the center of the radon band, but at its left edge. The phase shift introduced by the MCSA is  $\pi(A-B)$  in bin 1 and  $\pi(D-E)$  in bin 2. Since  $D = A + \dot{f}$  and  $E = B$ , this is also  $\pi(A-B + \dot{f})$ , so again bin 1 and bin 2 differ only by  $\pi\dot{f}$ . Thus it can be seen that the phase shift between adjacent integration periods does not depend on the position of the signal within the radon band. Furthermore the phase change from one integration period to the next is constant, which means the phase changes linearly with time. Since generalized coherence is invariant under such phase changes, this shift has no effect and can be ignored.

Considering now the MCSA phase shifts between adjacent frequency bins in the same integration interval. In Figure A3.10.12 the lines have the same meanings as in Figure A3.10.11. Points A, B, and C are the midpoints of bins 1, 2, and 3 respectively. Point D is the intersection of the center of the radon band and the center time line. The radon transform requires a weighted average of all bins lying within the radon band. The amplitudes of the weighting factor are the respective areas within the band, as discussed elsewhere. The relative phases must also be taken into account for proper combining of the bins.

The position of the signal within the band makes no difference as was shown above, but some point must be chosen as a phase origin, so that all the bins can be adjusted to have the same phase. The logical choice for this is point D. The MCSA phase shift is  $\pi(A-D)$  for bin 1,  $\pi(B-D)$  for bin 2 and  $\pi(C-D)$  for bin 3. Thus the necessary correction factor is the negative of these, and would be applied in the form

$$z_i = c_1 z_A e^{j\pi(D-A)} + c_2 z_B e^{j\pi(D-B)} + c_3 z_C e^{j\pi(D-C)} \quad (A3.10-12)$$

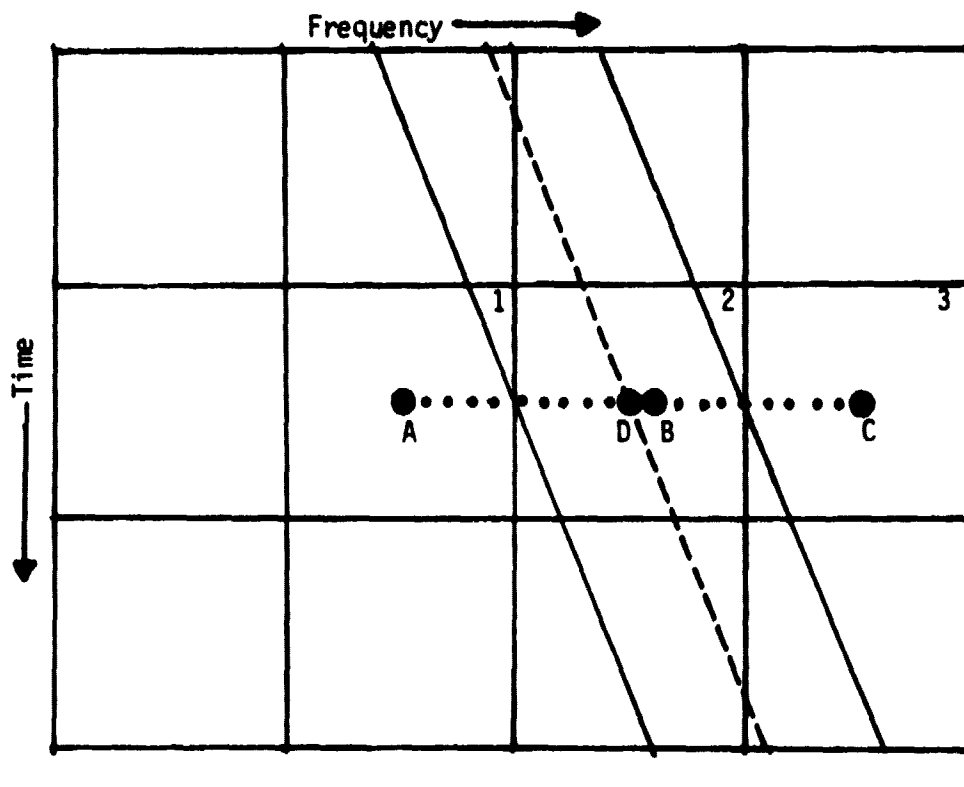


Figure A3.10.12 - MCSA Phase Shifts Between Adjacent Frequency Bins in the Same Integration Interval

where the c's are the amplitude weighting factors. It is also useful to note that  $(D-A) - (D-B) = B-A = 1$  and  $(D-C) - (D-B) = B-C = -1$  (since the units are bin-widths). Thus  $z_i$  can also be written as

$$\begin{aligned} z_i &= e^{j\pi(D-B)} (c_1 z_A e^{+j\pi} + c_2 z_B + c_3 z_C e^{-j\pi}) \\ z_i &= e^{j\pi(D-B)} (-c_1 z_A + c_2 z_B - c_3 z_C) \end{aligned} \quad (A3.10-13)$$

Either method (A3.10-12) or (A3.10-13) may be used to apply the necessary phase correction factors.

The two MCSA phase corrections derived above may be combined and applied simultaneously. The composite result is

$$z_i = c_1 z_A e^{j\pi(D-A-2fi)} + c_2 z_B e^{j\pi(D-B-2fi)} + c_3 z_C e^{j\pi(D-C-2fi)} \quad (A3.10-14)$$

for a given drift rate and starting point, the phase factors in equation (A3.10-14) may all be computed in advance, and stored in a lookup table for real time computation purposes.

#### Approximate Effect of Drift Rate Quantization

Since only a finite number of drift rates are used in the CWD it is important to know what happens if a signal drifts at rate other than one of those chosen for analysis. The worst case is when the signal drift rate is halfway between two to those chosen. That case will now be examined.

The CWD rate quantization chosen is such that the drift increment is the reciprocal of the total integration length (e.g., for 20 seconds integration, the drift increment is 1/20 Hz).

A typical worst case is thus a signal with a drift rate of 1/40 Hz. Over the total integration period, it will drift half-way across a radon band.

Then the worst way in which it can do so is to drift from the center to one edge. Since the combining operation among bins at the same time does not depend on the signal frequency, the combining will still be correct. However, the phase correction of the sum to the center of the radon band will have an error varying linearly from 0 at the beginning to  $\pi/2$  at the end of the integration period. Similarly the phase correction factor due to the signal itself will be in error ranging linearly from 0 to  $\pi$ , over the same period, resulting in a total error of  $3\pi/2$ . The net result is that the  $z_{i+1}z_i^*$  terms of GCV do not all add up exactly in phase, and partial cancellation occurs. The amplitude of the observed signal also changes as it drifts across the MCSA sinc response function. The situation is exactly analogous to that given in Table A.2, if it had an entry for  $0 \leq d \leq 0.75$ . Plotting and interpolating Table A.2 gives an amplitude value of 0.62 for this case. Thus the worst case signal loss is 2.1 db.

For the same drift rate, the best case would be for the signal to drift from  $-1/4$  bin, to  $+1/4$  bin, relative to the center. Then the amplitude factor would be about .85, making the average be about 0.74, or 1.3 db loss. If we now calculate the average over all drift rates, we can estimate the average loss to be one half of this or about 0.7 db.

In summary, by choosing the drift rate increment to be the reciprocal of the integration time, we incur an average decrease in sensitivity of 0.7 db, and worst case loss of 2.1 db. This loss is not substantial, and could be decreased by increasing the number of drift rates analyzed.

### Approximate Detection Statistics

In this section, approximate detection statistics for the CWD will be derived. The basis for this derivation is the body of pre-existing knowledge concerning the Radon transform. The derivation is approximate in the sense that it does not explicitly include the effects of Generalized Coherence. Based on the results of Section 3.3 above, these results may be considered as underestimates of the true performance of the CWD.

The particular angle-radius parametrization (which comes out of Radon transform theory naturally) utilized by the CWD was not developed for the Hough transform until (Duda and Hart, 1972) pointed out the advantages in 1972. Since this work many papers on the Hough transform and more general Hough-type transforms have appeared (Shapiro, 1978; Sklansky, 1978; Shapiro and Iannino, 1979). (Shapiro, 1978) and (Shapiro and Iannino, 1979) have studied the Hough transform in the presence of noise with particular attention given to errors in parameter space which arise from errors in feature space. Noise analysis of the MCSA output is discussed elsewhere in this report. The important point here is that the transform which is being used has been applied to arrays of data which are in many respects similar to the properties of the GCV and power spectrum data. The noise properties of the transform have been studied and it may be comforting to observe that the transform is in some sense optimal in that it behaves as a matched filter (Sklansky, 1978). Noise analyses from a slightly different viewpoint (Hanson, 1979) demonstrate that once the projections (GCV numbers or sum over power spectra along rays of width  $\underline{a}$ ) are obtained for a fixed angle  $\theta$ , the

function  $P_0(x)$  which contains both signal and noise information may be studied by standard techniques of communication theory (Middleton 1960; Middleton 1965; Whalen 1971). It follows that the optimum SNR is given by

$$\text{SNR} = \frac{\frac{1}{L} \left( \int_0^L |P_0(x)|^2 dx \right)^{1/2}}{\sqrt{a} \sigma / \sqrt{L}}$$

for a projection ray of width  $a$  and length  $L$ , for white noise with mean zero and variance  $\sigma^2$ , where the integral is over the ray width. Observe that for signal  $S$  in  $k$  pixels of unit width this reduces to  $kS/\sigma\sqrt{L}$ . For  $L = 1$  and  $k = 1$  we obtain  $S/\sigma$  for a single pixel, and for  $L = k$  we obtain  $\sqrt{k} S/\sigma$ .

Assume that  $L$  is the integration period  $t$  of the CWD. Then, according to the above SNR formula the optimum signal power extracted is given by  $\frac{k}{t} \sqrt{t} \frac{S}{\sigma}$ . This expression can be used to do some elementary analysis to investigate the set of conditions under which a signal can be detected if the angle of the drifting signal is not matched exactly by a projection ray  $P_0$  of the detector. The probability of detecting a signal can be analyzed with respect to the fraction of signal ( $k/t$ ) picked up by the detector at various false alarm rates ( $\alpha$  levels) and for different signal to noise ratios ( $S/\sigma$ ). In addition, the number of seconds needed for the integration period  $t$  in order to achieve a 99% probability of detection for different  $\alpha$  levels and  $S/\sigma$  is considered.

The results given below are based on the following: The noise is assumed to be a Gaussian (normally distributed) variable with mean power  $\mu/\text{Hz}$  and standard deviation  $\sigma = \mu$ . (Note that we do not assume normalization with zero mean at this stage.) The signal is assumed to be continuously drifting, with a constant mean power  $S/\text{Hz}$  and, if present, is presumed to exist for the entire integration period ( $t$  seconds). It is

also assumed that the average power  $\bar{x}$  (or GCV) in an arbitrary ray (traversing a sufficient number of data points) is nearly normally distributed with mean  $\mu$  and variance  $\sigma^2/t$  where  $t$ , the integration period, and  $a$  the width of the ray, are the dimensions of the ray. The standard deviation is, therefore,  $\sigma_{\bar{x}} = \sigma/\sqrt{t}$ .

The hypotheses to be tested are:

$H_0$  = The power distribution (in the region) is due to noise alone.

$H_1$  = There is a signal in addition to the noise.

The decision rule will be to reject  $H_0$  if  $\bar{x} > \mu + z_{\alpha} \sigma_{\bar{x}}$  where  $z_{\alpha}$  is the z-score associated with a false alarm rate  $\alpha$  (i.e.,  $\Pr(z > z_{\alpha}) = \alpha$ ). Therefore, the probability of detecting a signal when it is present is, in general,  $\Pr(\bar{x} > \mu + z_{\alpha} \sigma_{\bar{x}})$ .

Now, under  $H_1$ , there is signal power in addition to noise so the expected value of  $\bar{x}$  is  $\bar{x} = \mu + \frac{kS}{ta}$  where  $k$  is the number of times the signal is in the ray during the integration period. Since the signal power is assumed to be a constant,  $\bar{x}$  is the same value as before (i.e.,  $\sigma_{\bar{x}} = \sigma/\sqrt{ta}$ ).

In the special case of the CWD a projection ray is approximately 1 Hz wide ( $a = 1$ ) and  $t$  seconds long, so the value of  $\bar{x}$  is  $\mu + \frac{kS}{t}$  where  $\frac{k}{t}$  is that part of the time a signal is picked up by a projection ray.  $\mu$  is the mean noise power,  $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{t}}$ , and the signal to noise ratio is  $\frac{S}{\mu}$  (the average signal power divided by the average noise power).

By definition a z-score is  $z = \frac{y - \bar{y}}{\sigma_{\bar{y}}}$  where  $\bar{y}$  is the mean,  $y$  a specific value and  $\sigma_{\bar{y}}$  is the standard deviation. In this case it is given by  $z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$ . Then:  $1 - \beta$  = probability of detection =  $P_{\text{det}} =$

$$\begin{aligned}
P_{\text{det}} &= \Pr(\bar{x} > \mu + z_{\alpha} \sigma_{\bar{x}}) = \Pr\left( \frac{z > \mu + z_{\alpha} \frac{\sigma}{\sqrt{t}} - (\mu + \frac{kS}{t})}{\sigma_{\bar{x}}} \right) \\
&= \Pr\left( z > z_{\alpha} - \frac{S}{\mu} \left( \frac{k}{\sqrt{t}} \right) \right) \\
&= 1 - \Pr\left( z - \frac{S}{\mu} \frac{k}{\sqrt{t}} \sqrt{t} \right)
\end{aligned}
\tag{A3.10-15}$$

Table A3.10.3 gives the number of seconds needed in a time block to achieve a 99% probability of detection for various  $\alpha$ ,  $k/t$ , and  $S/\mu$ . Figures A3.10.13 (a), (b), and (c), and Tables A3.10.4 (a), (b), and (c), show the probability of detection versus time for various  $k/t$ ,  $P/u$ , and  $\alpha$  levels using equation (A3.10-15).

Based on these statistics, the probability of detecting a signal after 20 seconds of integration when  $\text{SNR} = 1$ , the false alarm rate  $\alpha = .01$ , and the angle of drift coincides exactly with the slope of some ray, is approximately 98%. If the drift angle is such that a projection ray picks up half the available signal, then the probability of detection for the same parameters specified above, degrades to 46%. There are several ways to increase the probability: one is to raise the false alarm rate, i.e. if  $\alpha = .05$ , then the probability of detection is increased by 72%; another is to increase the integration time, i.e., if  $t = 40$ , then when  $\alpha = .01$ , the probability of detection is about 85% and when  $\alpha = .05$  the probability of detection increases to 94%. Longer integration periods are possible by combining the approximate GCV numbers (ray sums) every 20 seconds.

It appears from the foregoing analysis that at least one projection ray must pick up half the signal, if present, in order to detect the signal. Therefore, the number of projection rays to include in the detection mask is a function of that constraint.



$\alpha = .25$

<u>S/N</u>	<u>k/t</u>	<u>time(sec.)</u>
.8	1/2	56
.8	1/3	127
.8	1/4	225

$\alpha = .05$

.8	3/4	44
.8	1/2	99
.8	1/3	223
.8	1/4	396
.5	3/4	113
.5	1/2	253
.5	1/3	570
.5	1/4	1014

$\alpha = .01$

1	1/2	87
.8	3/4	60
.8	1/2	136
.8	1/3	305
.8	1/4	543
.5	3/4	154
.5	1/2	347
.5	1/3	782
.5	1/4	1390

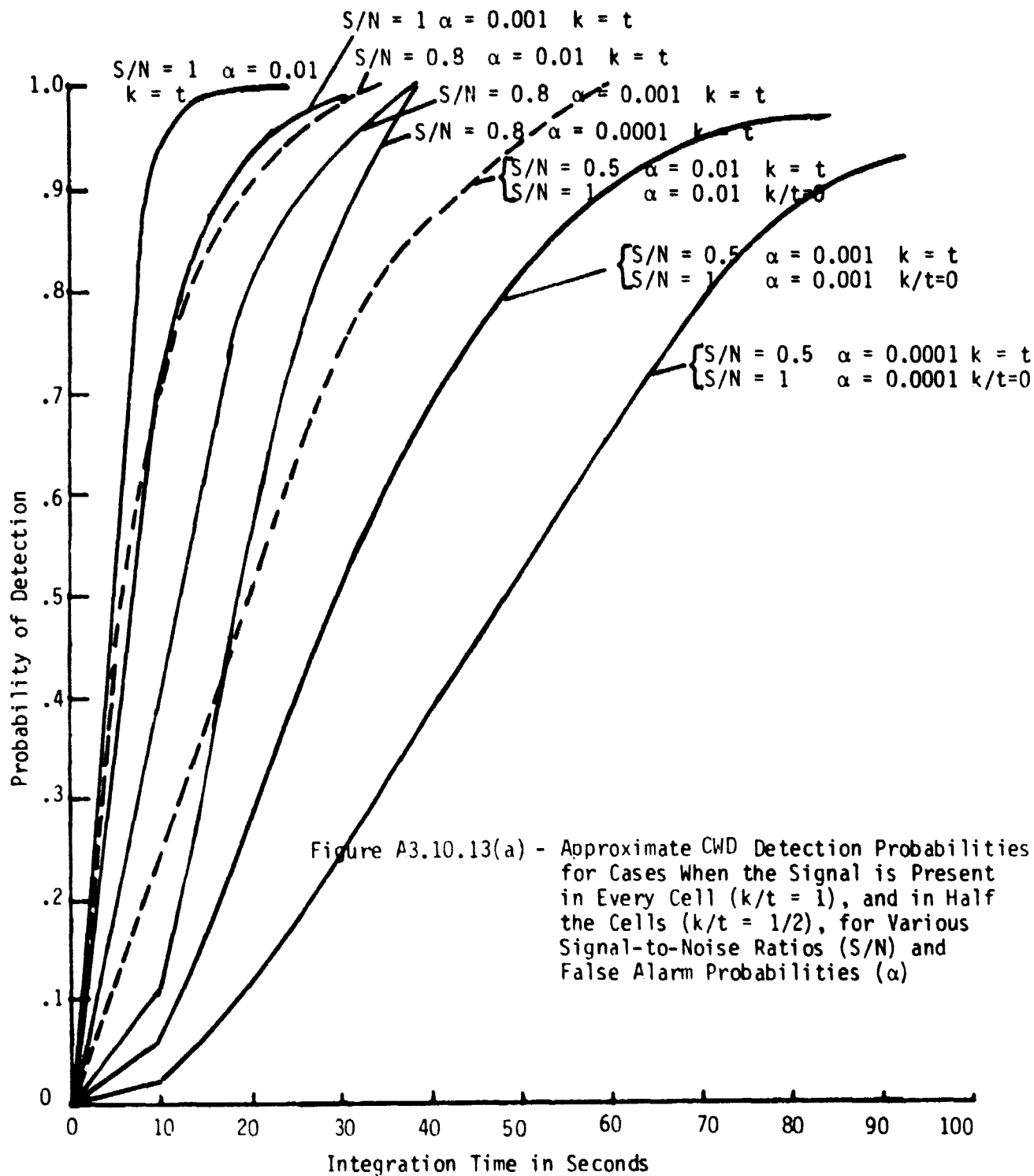
$\alpha = .001$

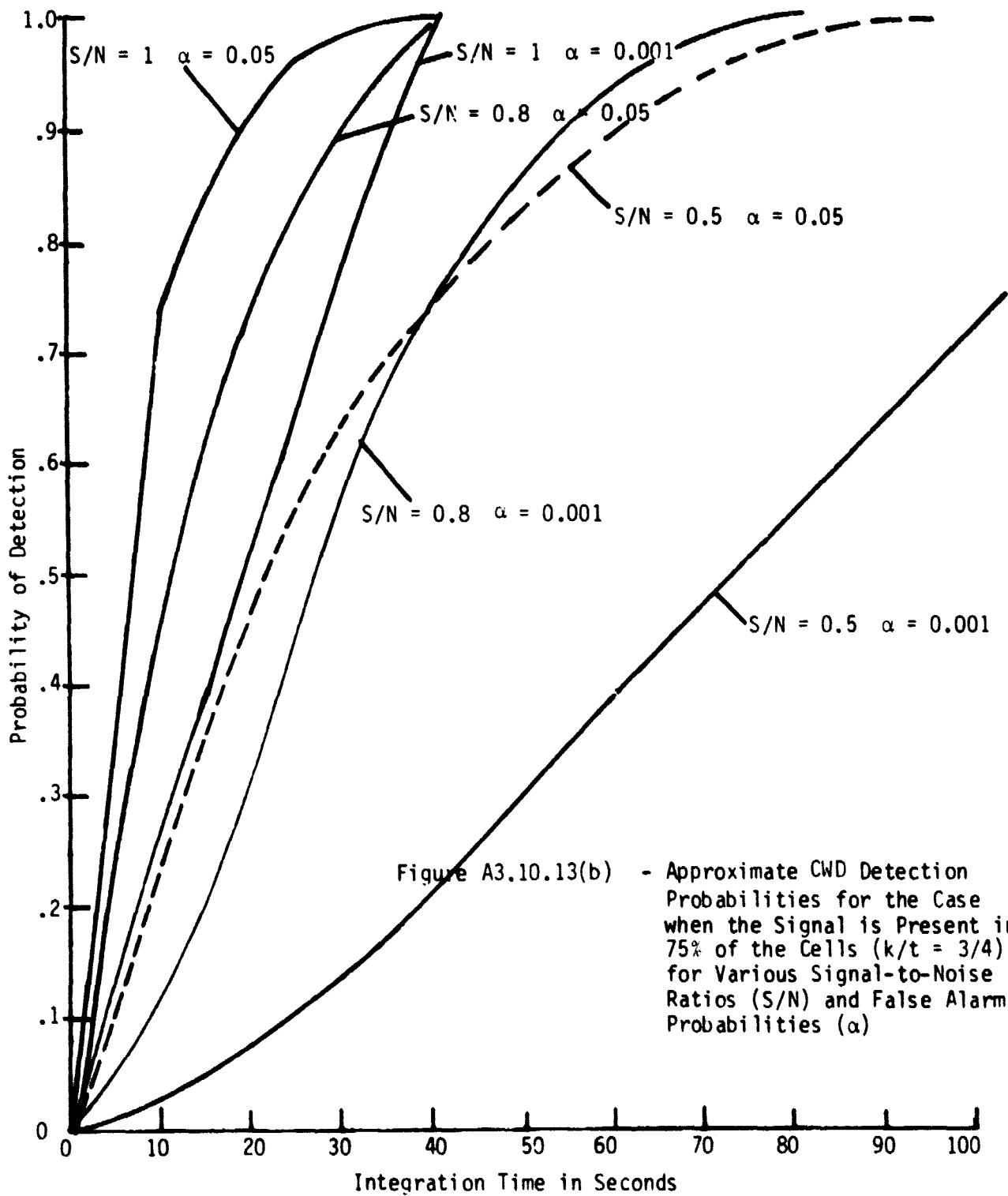
1	1/2	118
.8	3/4	82
.8	1/2	184
.8	1/3	410
.8	1/4	734
.5	3/4	209
.5	1/2	170
.5	1/3	1058
.5	1/4	1880

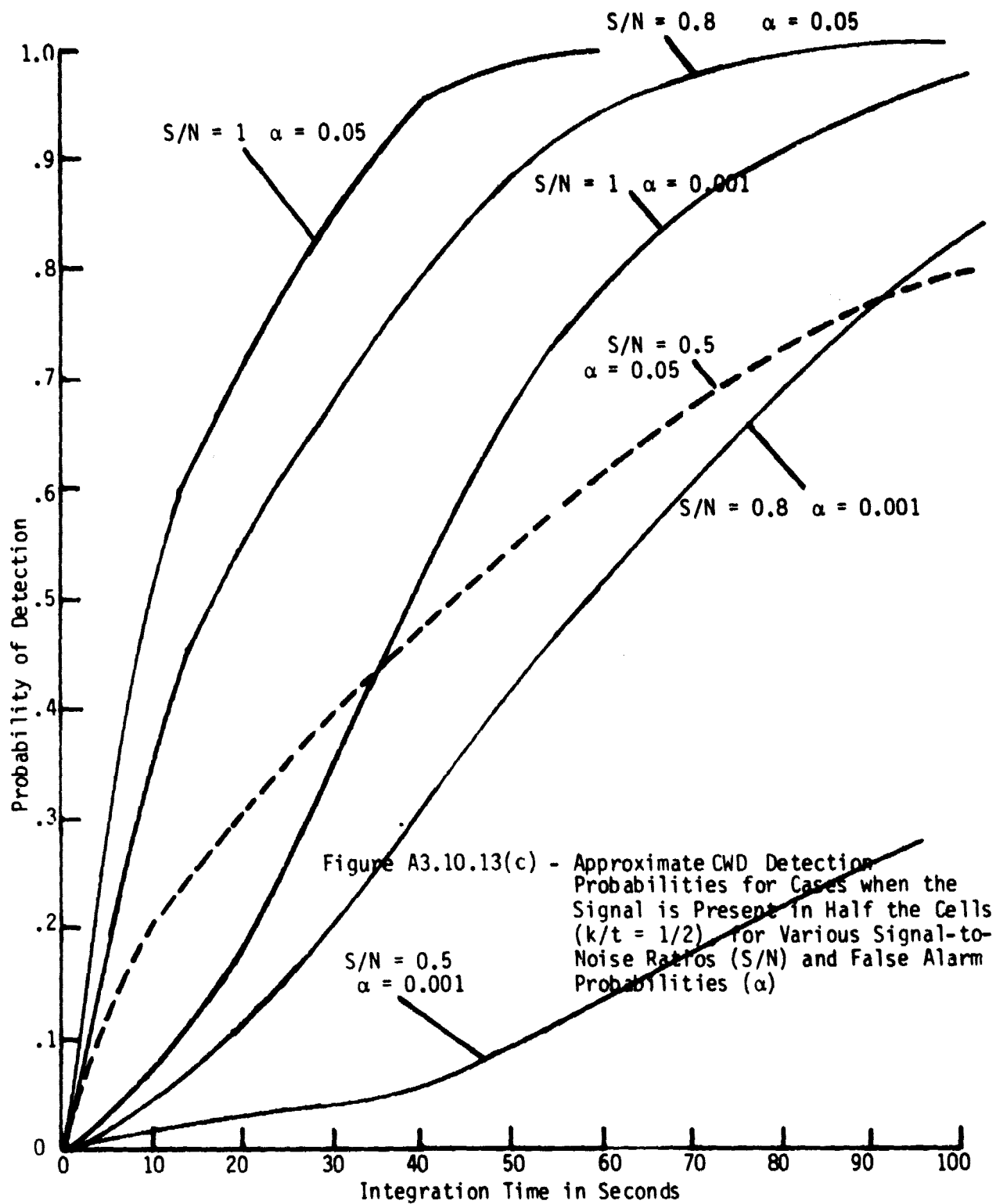
$\alpha \approx 10^{-7}$

.5	1	54
1	1/2	215
1	1/2	335
.8	1/2	860

Table A3.10.3 - Approximate Number of Seconds Needed for the Integration Period to obtain a 99% Probability of Detection at Various  $\alpha$  and S/N.







$\alpha \backslash t$	0	10	20	30	40	50	60	70	80	90	100
<u><math>S/\sigma = 1</math></u>											
.01	1%	80%	98%	99.9%							
.001	.1	53	92	99	99.9						
<u><math>S/\sigma = .8</math></u>											
.01		57	89	98	99.8						
.001	.1	29	69	90	98						
.0001	.01	12	45	75	91	97.5					
<u><math>S/\sigma = .5</math></u>											
.01		23	46	66	80	89			98		
.001		6	20	36	53		78			97	
.0001		1.7	7	17	29.5		57			90	
<u><math>S/\sigma = .3</math></u>											
.01		8.4	16	25	33						
.001		1.6	4	7.4	11.7						
.0001				1	2						

Table A3.10.4(a) - Approximate Probability of Detection for Various  $\alpha$  and Integration Periods  $t$ , when  $k/t = 1 =$  Perfect Match

$\alpha \backslash t$	0	10	20	30	40	50	60	70	80	90	100
<u><math>S/\sigma = 1</math></u>											
.05		76	96	99	100						
.01		52	85	96	100						
.001		24	60	85	100						
.0001		10	36	66	100						
<u><math>S/\sigma = .8</math></u>											
.05		60	85	95	100						
.01		33	64	83	93						
.001		12	32	58	76		94			99	
.0001		4	15	34	54						
<u><math>S/\sigma = .5</math></u>											
.05		32	51	66	76		89			98	
.01		13	26	39	52						
.001		3	8	13	24		43			75	
.0001		.6	2	5	9						

Table A3.10.14(b) - Approximate Probability of Detection for Various  $\alpha$  and Integration Periods  $t$ , when  $k/t = 3/4$

$\alpha \backslash t$	0	10	20	30	40	50	60	70	80	90	100
<u><math>S/\sigma = 1</math></u>											
.05	46	72	86	93	97						
.01	22	46	66	80							
.001	7	20	36	53		78				97	
.0001	2	7	17	30		57				90	
<u><math>S/\sigma = .8</math></u>											
.05	35	56	71	81		93				100	
.01	14	30	44	58							
.001	3	10	18	29		50				82	
.0001	.7	3	7								
<u><math>S/\sigma = .5</math></u>											
.05	20	30	39	47		61				80	
.01	6	11	17	23							
.001	1	2	4	7		12				28	
.0001	.2	.5	1	2							

Table A3.10.14(c) - Approximate Probability of Detection for Various  $\alpha$  and Integration Periods  $t$ , when  $k/t = 1/2$

Consider the situation depicted in Figure A3.10.14 where the signal lasts  $t$  seconds and falls within the ray labeled S which is a perfect match. Suppose the signal is traversed by a projection ray  $R_e$ ; then the amount of the signal picked up by the ray is proportional to the shaded area A and  $A = 1/2 \{(\cos(2\theta_s) + 1)\cot\Delta\theta - \sin(2\theta_s)\}$  where  $\theta_s$  is the angle the signal makes relative to the vertical and  $\Delta\theta$  is the difference between the ray angle  $\theta_R$  and the signal angle.

For the implementation of the algorithm the area of every ray is the same, no matter what the angle of the projection is. Also, that area of a ray of width 1 unit and length  $t$  seconds is  $t$  square units. Therefore, in order to ensure that half the signal falls within the a projection ray we need to have

$$t = (\cos(2\theta_s) + 1)\cot\Delta\theta - \sin(2\theta_s)$$

so

$$\cot\Delta\theta = \frac{t + \sin(2\theta_s)}{\cos(2\theta_s) + 1}$$

Now, as  $\theta_s \rightarrow 0$ ,  $\cot\Delta\theta \rightarrow t/2$ , and as  $\theta_s \rightarrow 45^\circ$ ,  $\cot\Delta\theta \rightarrow \frac{t+1}{1}$ .

If  $t = 20$  seconds,  $\Delta\theta$  is between  $3^\circ$  and  $6^\circ$  whenever  $\theta_s$  is between  $0$  and  $45^\circ$ .

Consequently, if the ray projections of the CWD mask are separated by  $3^\circ$  it should be possible to detect a signal with an acceptable probability. Furthermore, when  $t = 20$ ,  $\tan^{-1} \frac{1}{20} \approx 3^\circ$  which means that the ends of the rays overlap.

A second consideration involved in the choice of the number of angles is the maximum amount of Doppler shift caused by planetary rotation for a planet that could support life. For a planet like the earth, the Doppler shift is .15Hz/sec. (CYCLOPS, p.57); for Jupiter it is 14 Hz/sec, and for



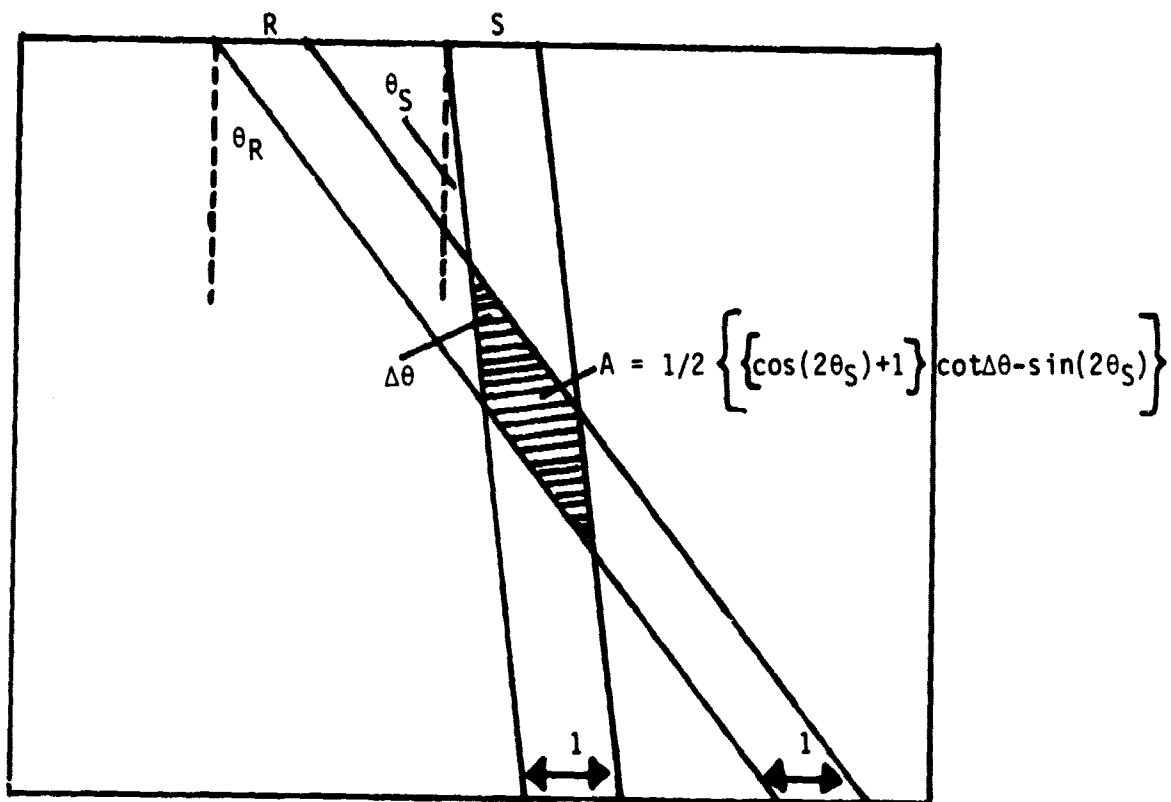


Figure A3.10.14- Proportion of Area in the Intersection of an Arbitrary Ray R with an Ideal Ray S, which Contains the Signal

an Earth-sized planet rotating at Jupiter's rate (8 hr) it is 1.4 Hz/sec. (Machol, 19 ).

Finally, there are implementation and cost considerations. A reasonable choice, for a first approximation, seems to be 20 angles spaced  $3^{\circ}$  apart covering the range from  $-27^{\circ}$  to  $+27^{\circ}$ .

#### Combination of Successive Generalized Coherence Values

The sets of generalized coherence values calculated every 20 seconds are used not only for immediate threshold testing, but are accumulated as time goes on to obtain additional results encompassing all the data received up to that time.

Both the old and new sets of accumulators each contain GCV's for each of eight million frequency channels for forty-one uniformly spaced drift rates between plus and minus one hertz per second. However, it is not optimal to simply combine corresponding frequency-drift-rate cells.

When two sets of accumulators are combined, those individual GCV's are added together which could approximate the same uniformly drifting signal. For example a typical low level signal might be present most strongly in accumulator A from one set and accumulator B in the next set. Most likely A and B would have the same drift rate and the reference frequency of B would be offset from A by the drift rate times 20 sec. Note that the correspondence between one set of accumulators and the next depends on their individual drift rates.

If a uniformly drifting signal had a drift rate that exactly matched one of the accumulators, the best way to combine the GCV's from the

previous time periods with the GCV's for the present 20 sec would be to add the values of corresponding accumulators together. This would always add accumulators containing the signal together and would always add accumulators containing only noise to other accumulators that contain only noise. This would allow signal-containing accumulators to be most clearly distinguished from those that do not.

However, a real signal could have any drift rate and would in general be expected to gain or lose frequency relative to the closest accumulator drift rate. In fact, from one set of accumulators to the next, a signal could go to the next higher or lower drift rate as well as the next higher or lower frequency.

An initial step in resolving this ambiguity is to offset the accumulator frequencies by one-half hertz in alternate 20 second sets. This reduces to four the number of accumulators in one set that could contain the same signal as a particular accumulator in the next 20 second set.

For a signal that is not clearly distinguished from noise one cannot be sure which of the four accumulators from the previous set might have contained the signal. The best estimate is to choose the one with the highest GCV. Conceptually, the procedure is to add the maximum of four possibly corresponding accumulators from one set to each of the accumulators in the next set. This modified set of accumulators is then used in the same way to modify the following set and so on for as many sets as desired.

Every accumulation produced from a combination of 20 second periods is tested against an appropriate threshold. Whenever a threshold is exceeded it is considered sufficient evidence that a signal is present and a flag is set for the control computer.

How often have I said to you that when you have eliminated the impossible, whatever remains, however improbable, must be the truth.

--AC Doyle, Sign of Four

### Appendix A3.11

#### Cluster Detector Details and Statistics

##### Step 1: The Histogram Analysis

We initially subdivide the 50 x 250,000 compacted array into 1250 pattern blocks and have dedicated processors operate on each pattern block independently. Subdividing the compacted array vastly reduces the number of computations and, hence, hardware requirements, and does not result in a significant loss of information for the second step analysis. After subdivision, a processor establishes a histogram for its pattern block by simply "counting" the number of pattern points in each pattern class. If the number of pattern points in some pattern class exceeds a preset threshold, the central computer is alerted. An excess number of points in a pattern class within a pattern block suggests the presence of a signal due to an unusual number of one or more NBIT tests scoring a pass ( = "1"). This possibility may be confirmed by the clustering algorithm of step 2, which detects the spatial structure of these pattern points within the pattern block. The rationale for counting the pattern points in the pattern classes instead of counting identical types of pattern points within a pattern block is provided by the following example.

Suppose that in some pattern block there is a narrow band signal which consistently sets the same three bits for a 1000 seconds. By our supposition, the signal will be present in 50 cells of the pattern block. A pattern point affected by the signal has six bits which remain unaffected. Each of these six bits has a 50% false alarm rate. Hence a pattern point affected by the signal has equal probability of belonging. Although the signal to noise ratio remains essentially the same, we have a clustering of  $1294 - 1250 = 44$  points in this class as compared to less than one for the other histogram. Furthermore, the three 2-bit and three 1-bit subclasses of the above 3-bit pattern class will also show significant clustering. This provides us with another indicator as to whether or not a pattern class is affected by a signal.

As was indicated in Section IIID4.5, the first stage clustering serves two basic functions. One is to compactify the data even further, in the sense that the input to the second step analysis consists of only those pattern points whose classes are deemed "interesting" by the first step. The second function of the first step clustering is to alert the central computer if some pattern class contains an unusually large number of pattern points. In order to achieve this objective we establish thresholds for the histogram such that whenever the number of pattern points in a particular pattern class exceeds its threshold, the alarm is activated. The levels of the thresholds are such that the false alarm rate will be within acceptable limits. We expect some readjustment of these thresholds as more data concerning the interdependency of the 9-bit tests becomes available during actual simulations.

Let  $i = 1, 2, \dots, 511$  identify the pattern class (we do not consider the zero or "no bits are on" pattern class). Let  $n(i) = 1, 2, \dots, 9$  the number of bits which determine the  $i$ -th class,  $N_i$  = the number of pattern points in the  $i$ -th class/pattern block, and  $T_{n(i)}$  the threshold for the  $i$ -th class. If we compare the desired false alarm rate with tossing a coin once every 1,000 seconds, where "heads" corresponds to "no false alarm" and "tails" to "one or more false alarms", and we want the coin to be honest, then we must set

$$\Pr(x > 0) = 1 - \Pr(x = 0) = 1 - (1 - \alpha)^{1250} = 0.5.$$

where  $\alpha = \Pr(N_i > T_{n(i)} \text{ for at least one } i)$   
 $= \Pr(\text{at least one threshold exceeded})$

and  $x \sim b(1250, \alpha)$ .

Solving for  $\alpha$ , we obtain

$$1 - \alpha = (0.5)^{1/1250} = 0.999445636 \text{ or}$$

$$\alpha = 0.000554364$$

We note that

$$\Pr(x = 1) = 1250\alpha(1 - \alpha)^{1249} = 0.347,$$

$$\Pr(x = 2) = \frac{1250 \times 1249}{2} \alpha^2(1 - \alpha)^{1248} = 0.1201 \text{ and}$$

$$\Pr(x = 3) = \frac{1250 \times 1249 \times 1248}{6} \alpha^3(1 - \alpha)^{1247} = 0.0277.$$

Thus the probability of two or three false alarms in 1,000 seconds is very small. If  $\alpha_i = \Pr(n_i > T_{n(i)})$ , then  $\Pr(N_i \leq T_{n(i)}) = 1 - \alpha_i$ .

Since  $\Pr(N_i \leq T_{n(i)} \forall i = 1, 2, \dots, 511) = 1 - \alpha$ , we have

$$\prod_{i=1}^{511} (1 - \alpha_i) = (1 - \bar{\alpha})^{511} = 1 - \alpha \quad \text{or}$$

$$(1 - \alpha) = (0.999445636)^{1/511} = 0.9999989148.$$

The corresponding normal deviate is  $z_{\bar{\alpha}} \approx 5$ . Therefore,

$$\frac{T_{n(i)} - \mu_{n(i)}}{\sigma_{n(i)}} = 5 \quad \text{or}$$

$$T_{n(i)} = 5\sigma_{n(i)} + \mu_{n(i)}, \quad \text{where}$$

$\mu_{n(i)}$  denotes the expected number of pattern points in the  $i$ -th class under noise alone and  $\sigma_{n(i)}$  denotes the corresponding standard deviation.

We define the sensitivity,  $S_{n(i)}$ , associated with the threshold  $T_{n(i)}$  as the minimum number of cells/pattern block a signal must occupy in order for  $\bar{\mu}_{n(i)} > T_{n(i)}$ , where  $\bar{\mu}_{n(i)}$  denotes the expected value of the signal in random noise. If the signal occupies  $k$  cells, then from the inequality

$$\bar{\mu}_{n(i)} = \frac{10,000 - k}{2^{n(i)}} + k > T_{n(i)}$$

we obtain

$$k > \frac{2^{n(i)} T_{n(i)} - 10,000}{2^{n(i)} - 1}$$

Hence,

$$S_{n(i)} = \left\lceil \frac{2^{n(i)} T_{n(i)} - 10,000}{2^{n(i)} - 1} \right\rceil, \quad \text{where}$$

$\lceil x \rceil$  denotes the first integer greater than  $x$ .

For  $\alpha = 0.000554364$  and  $z_{\bar{\alpha}} = 5$ , we obtain the threshold sensitivities shown in Table A3.11.1.

$n(i)$	$\sigma_{n(i)}$	$h_n(i)$	$T_n(i)$	$S_n(i)$
9	4.42	19.53	42	23
8	6.32	39.06	71	33
7	8.8	78.13	122	43
6	12.4	156.25	218	62
5	17.4	312.5	400	91
4	24.216	625	746	130
3	33.07	1250	1415	189
2	43.3	2500	2717	290
1	50	5000	5250	500

Table A3.11.1 - Threshold Sensitivities



## Step 2: Single Linkage Cluster Analysis

This algorithm searches for spatially contiguous cells of a pattern type within a pattern block. Two cells are contiguous with respect to the  $i$ -th pattern class if they are within a pre-specified distance  $d_i$ . If contiguous cells form a group such that the number of cells within the group exceeds a preset threshold  $T_i$  depending on the  $i$ -th pattern class, then a cluster exists and the central computer is alerted.

The input to the algorithm are the pattern points of pattern classes deemed interesting by the histogram, where "interesting" is defined as follows. Let  $j = 1, 2, \dots, 1250$  identify the pattern block,  $N_{ij}$  = the number of pattern points in the  $i$ -th class and  $j$ -th pattern block, and set  $z_{ij} = \frac{N_{ij} - \mu_n(i)}{\sigma_n(i)}$ . The interesting classes and pattern blocks are those which have large normal deviations  $z_{ij}$ . Because of the compactification scheme, we have the real-time ability for approximately 100 searches/10 seconds. Hence we may search for clusters in pattern blocks producing the 100 largest normal deviations.

The implementation of cluster seeking is based on the general algorithm described next. The algorithm will vary somewhat according as to whether we are considering one and two bit pattern classes, or three and greater bit classes. For each large  $z_{ij}$ , the algorithm establishes three lists,  $A_{ij}$ ,  $B_{ij}$ , and  $C_{ij}$ . Initially  $A_{ij}$  is a list of cell positions of all cells in the  $j$ -th pattern block which are occupied by pattern points of the  $i$ -th class;  $B_{ij}$  is a list of distances between all pairs of occupied cells sorted by ascending distance with ties arranged arbitrarily, and  $C_{ij} = 0$ . The algorithm proceeds in three steps.

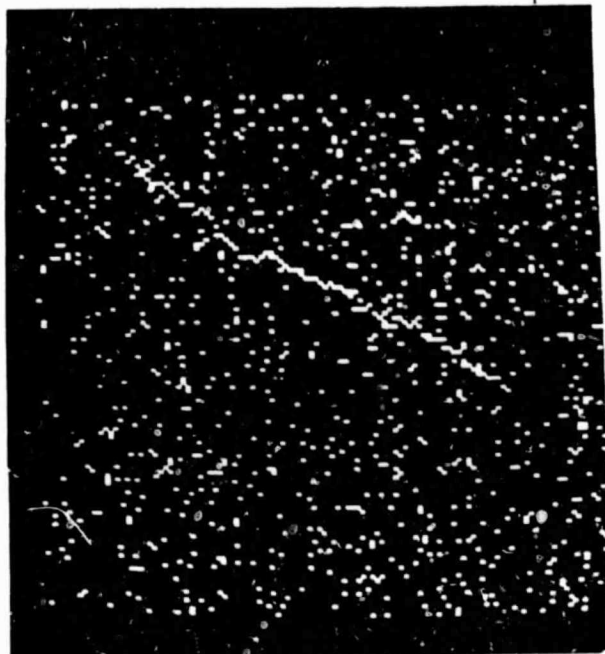
1. If the first distance in  $B_{ij}$  is greater than  $d_i$ , stop.  
There is no cluster. Otherwise, remove the first entry from  $B_{ij}$ , and remove the cells corresponding to this distance from  $A_{ij}$  and add them to  $C_i$ .
2. Update the list  $B_{ij}$  by computing the distances from cells in  $A_{ij}$  to the "cell"  $C_{ij}$ .
3. If every distance to  $C_{ij}$  is greater than  $d_i$ , return  $C_{ij}$  to its original state and go to 1. Otherwise remove from  $B_{ij}$  all distances between cells of  $A_{ij}$  and the cluster  $C_{ij}$  which are less than or equal to  $d_i$ . Remove the cells corresponding to these distances from  $A_{ij}$  and add them to  $C_{ij}$ . If the number of pattern points in  $C_{ij}$  exceeds the threshold  $T_i$ , stop. A cluster is found. Otherwise go to 2.

---

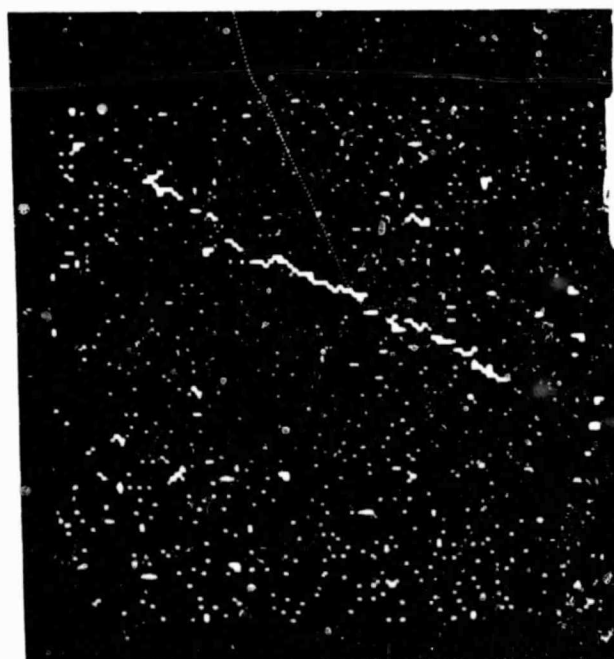
Note: For  $a \in A_{ij}$ ,  $d(a, C_{ij}) = \min \{d(a, c), c \in C_{ij}\}$

We may obviously continue the algorithm in order to establish the actual size of  $C_{ij}$  and/or for finding further clusters in the remaining list  $A_{ij}$ . The actual implementation of the algorithm for three and larger bit classes is reduced to a simple scan and record procedure. A processor scans  $A_{ij}$  and if a cell of  $A_{ij}$  does not have a preset number  $N_i$  of neighbors within the distance  $d_i$ , the cell is removed from  $A_{ij}$ . After all cells which do not have  $N_i$  neighbors within  $d_i$  have been removed, the clusters remaining are then examined to see whether the number of cells in any cluster exceeds  $T_i$ . We may also repeat the entire procedure on the clusters obtained by redefining  $N_i$ . Figure A3.11.1(a) shows the 2-dimensional

array representing  $A_{ij}$ , where the  $i^{\text{th}}$  class is a 3-bit pattern class, before the application of the algorithm, while Figure A3.11.1(b) shows the result of applying the algorithm with  $d_i = 1$  and  $N_i = 2$ .



(a) - Before



(b) - After

Figure A3.11.1 - Effect of the Single Linkage  
Cluster Analysis Algorithm

The algorithm is somewhat modified for the one and two bit pattern classes. Since each bit has a 50% false alarm rate, half of all cells in a pattern block are expected to be occupied by any given 1-bit pattern class due to noise alone. This situation is illustrated in Figure A3.11.2 where the "on" cells in the pattern block belong to a particular 1-bit pattern class.

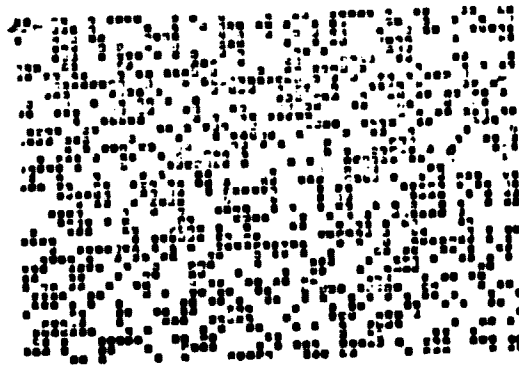


Figure A3.11.2 - A Pattern Block for the One Bit Pattern Class

It becomes clear from Figure A3.11.2 that there is an extremely high degree of contiguousness between the "on" cells. In fact, one can find connected sequences of "on" cells forming meandering lines which stretch across the entire pattern block.

Because of this high degree of contiguousness, we are forced to set  $d_i = 1$  for one and two bit classes and to alter the clustering algorithms in order to keep the false alarm rate within a reasonable minimum.

The altered algorithm is capable of finding clusters of cells containing horizontal and/or vertical line segments formed by contiguous cells, and of finding clusters containing rectangles of contiguous cells. For lines, there are two minimal alarm thresholds,  $T_1$  and  $T_2$ , with  $T_1 < T_2$ , while for clusters containing rectangles there is only one, which we shall

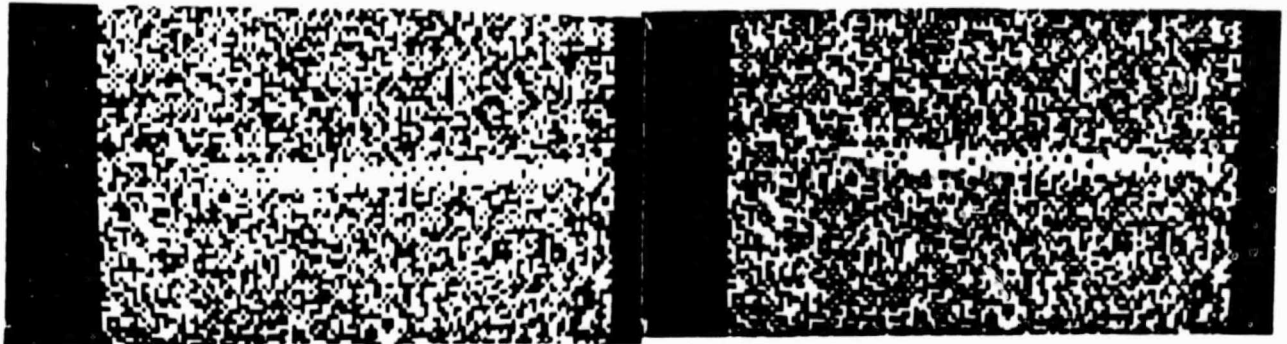
denote by  $T$ . We set  $T_2$  such that if the length of a contiguous sequence of cells in a row or column exceeds  $T_2$ , the probability of an alarm due to noise alone is extremely low. If  $T_2$  is exceeded, we say that a cluster has been found and alert the central computer. The algorithm itself is reduced to a scanning and recording procedure and is accomplished in the following steps.

1. Scan the first row of  $A_{ij}$ . If a contiguous sequence of "on" cells of length greater or equal to  $T_1$  is found add these cells to  $C_{ij}$  while removing them from  $A_{ij}$ . Also remove all cells from  $A_{ij}$  appearing before this contiguous sequence. If the number of cells in  $C_{ij}$  exceeds  $T_2$ , stop. A cluster has been found. If there is no sequence exceeding  $T_1$ , remove all cells in this row from  $A_{ij}$ .
2. Scan the next row. If another sequence of contiguous cells is found which exceeds  $T_1$  and is situated "below"  $C_{ij}$ , that is, the vertical projection of  $C_{ij}$  into the row below is contained in the sequence or the vertical projection of the sequence into the row above it is contained in  $C_{ij}$ , remove the cells from  $A_{ij}$  and add them to  $C_{ij}$ . Otherwise, return  $C_{ij}$  to its original state and go to 1.
3. If the number of cells in  $C_{ij}$  exceeds  $T$ , stop. A cluster has been found. Otherwise go to step 2.

If no clusters are found and  $A_{ij}$  is erased, we repeat the algorithm by refreshing  $A_{ij}$  with its original data, and by replacing "row" with "column". That is, we scan along columns. Obviously, the algorithm stops whenever all the data from  $A_{ij}$  has been exhausted. The effect of the algorithm is shown in Figure A3.11.3. Here (a) represents a pattern block

ORIGINAL PAGE  
BLACK AND WHITE PHOTOGRAPH

with "on" cells belonging to a specific 1-bit pattern class before application of the algorithm and (b) shows the cells remaining after application of the algorithm.



(a) - Before

(b) - After

Figure A3.11.3 - Effect of the Modified Single Linkage Cluster Analysis Algorithm for a One-Bit Pattern Class

Our final goal is to obtain a method for establishing some of the thresholds discussed above. Since we are interested in clusters formed by contiguous cells, the first question which arises is as to the probability, due to noise alone, of obtaining at least  $r$  contiguous ( $d_i = 1$ ) cells belonging to the same class in a row (column) of 200 (50) cells long. Calling a sequence of  $r$  contiguous cells in a row of  $n$  cells a run of  $r$  successes, we are asking for the probability of having a run of  $r$  successes in  $n$  trials, where  $n = 200$  or  $n = 50$ .

Let  $P_i$  denote the probability of the event that a cell contains a pattern point of the  $i$ -th class due to noise, and let  $P_{n,r}$  denote the probability of a run of  $r$  successes in  $n$  trials. Then  $P_{n,r}$  is given by

$$P_{n,r} = 1 - X_{n,r} + P_i^r X_{n-r,r} \quad ,$$

where

$$X_{n,r} = \sum_{k=0}^m (-1)^k \frac{(n - kr)!}{k! (n - kr - k)!} ((1 - P_i) P_i^r)^k$$

and  $m$  is the integer truncation of  $\frac{n}{r+1}$ . Values of  $P_{n,r}$  for  $n = 200$  are given in Table A3.11.1. We computed  $P_{n,r}$  for as many as fifty contiguous cells, an event which would certainly trigger alarms set in previous stages and in step 1 clustering. Values for  $T_1$  and  $T_2$  can be chosen from the  $r$  column of Table A3.11.1 with the corresponding  $P_{n,r}$  providing the false alarm rate. A rough estimate for  $T$  can also be computed with the aid of Table A3.11.1. Thus, if we choose  $T_1 = 8$  for a 1-bit pattern class and multiply  $P_{n,8}$   $k$  times, we obtain a rough estimate as to the probability due to noise, of obtaining a  $T = 8 \times k$  rectangle of contiguous cells.

Table A3.11.1 - Probability that a contiguous string of  $r$  cells, contained in a linear string of  $n$  cells, will exceed the detection threshold due to noise alone. Three examples are shown.

$P_i = 0.03125$		
$n$	$r$	$P_{n,r}$
200	2	1.7E-01
200	3	5.8E-03
200	4	1.8E-04
200	5	5.6E-06
200	6	1.7E-07
200	7	5.4E-09
200	8	1.7E-10
200	9	1.0E-11
200	10	8.8E-16

$P_i = 0.00781250$		
$n$	$r$	$P_{n,r}$
200	2	1.1E-02
200	3	9.3E-05
200	4	7.2E-07
200	5	5.6E-09
200	6	4.0E-11
200	7	1.7E-15
200	8	1.3E-17
200	9	1.0E-19
200	10	8.4E-22

$P_i = 0.00390625$		
$n$	$r$	$P_{n,r}$
200	2	3.0E-03
200	3	1.1E-05
200	4	4.5E-08
200	5	1.8E-10
200	6	3.5E-15
200	7	1.3E-17
200	8	5.4E-20
200	9	2.1E-22
200	10	8.2E-25



## Appendix A3.12

### An Evaluation of Some Classical Algorithms in the SETI Context

#### 1. Matched Filters

A possible SETI signal detection system might be constructed utilizing an array of matched filters. A matched filter is an algorithm for the combination of information from the MCSA in such a manner as to maximize, for a given expected set of signal characteristics, the peak power to root-mean-square (rms) noise power.

It can be shown (Schwartz, Bennet, and Stein, 1966) that if a signal, represented as a function in time by  $f(t)$ , passes through a linear filter, whose transfer function is  $H(\omega)$ , that filter is said to be matched if:

$$H(\omega) = F^*(\omega) e^{-j\omega t_0}$$

where  $F^*(\omega)$  = complex conjugate of the Fourier transform of  $f(t)$ ,

$t_0$  = time of occurrence of the peak value of  $f(t)$

This filter performs the function of:

- a) weighing the spectral components in the filter in proportion to the spectral components in the signal;
- b) aligning all spectral components in the filter in phase.

This is a type of conjugate matching.

If such a filter can be implemented, the ratio of the peak signal power to rms noise power, assuming white noise, is

$$\frac{S}{N} = \frac{2E}{KT}$$

where  $E$  = the total signal energy

$K$  = Boltzmann's constant

$T$  = system noise temperature

Thus, the detectability of a signal through a matched filter is independent of the form of the signal.

It is possible to impress information onto an otherwise monochromatic continuous carrier in a number of ways: amplitude, phase, frequency, and polarization. Terrestrial technologies have used all four possible modulations to some extent in order to convey messages via the electromagnetic radiation. One might expect that it could be a straightforward exercise to implement a large set of filters, a group for each modulation scheme, which would respond, in some optimal way, to extraterrestrial modulated RF energy. In order to accomplish this, however, one must know a priori, the form of the message contained within the modulation. Since this is quite obviously impossible, this direction for a SETI detector is not reasonable.

We expect that an ETI signal, deliberately radiated for the purposes of detection (a beacon) by other civilizations, would be a "single" signal in order to increase the likelihood of detection (Dixon, 1973). Simplicity, in terms of detectability, implies extreme spectral purity (Seeger, 1971; Oliver and Billingham, 1971). Further, many forms of terrestrial modulation use the redundant information contained within the monochromatic carrier to facilitate coherent detection. Therefore, the strategy is to attempt to detect the narrowband carrier, and to leave the examination of possible modulation sidebands, however they may manifest themselves, to a purely statistical approach. The advantage of this latter point is that we search the MCSA output only for possible departures from noise statistics.

Signals which are pulsed in a time scale shorter than the reciprocal bandwidth of the MCSA (1 Hz) present special problems to the signal detector.

The present design of the MCSA precludes optimum detection of pulses between time frames from 1 second to  $\geq 1/8000$  second, and from  $\leq 1/8000$  to  $\sim 10^{-9}$  second, the reciprocal IF bandwidth. Pulsed signals in those realms must be found by an analysis of the time domain signal, rather than the frequency domain signal as presently envisaged in the MCSA.

## 2. Walsh Transforms

As a cartoonist captures the important features which make a simply-drawn caricature recognizable, it would be desirable to have a transform which would extract --from the MCSA output --recognizable features necessary to indicate the presence of a signal. Walsh functions and transforms were explored as possible candidates for that transform.

The Walsh functions may be defined as an orthonormal set of square wave functions.

The Walsh function has two arguments:  $n$  is the order and  $u$  is normalized time. The order  $n$  is related to the number of zero crossings in the interval over which the function is defined. (See Fig. A3.12.1 taken from Harmuth, 1977).

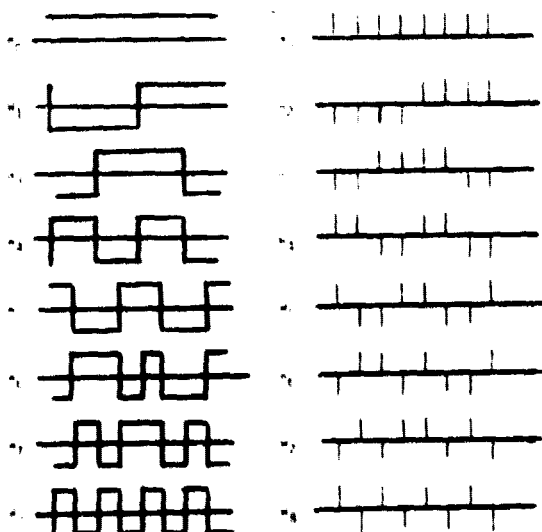


Figure A3.12.1 - Walsh Functions, Continuous and Discrete

Each Walsh function may be expressed as a product of odd symmetry periodic square waves of amplitude  $\pm 1$ . (Rademacher functions)

Walsh functions may be transformed into the "sequency" domain, where sequency is the number of zero crossings. A sequency spectrum plot would indicate the number and magnitude of the coefficients necessary for a "Walsh series" expansion. (See Figure A3.12.2 taken from Alexandridis, 1971) (Schreiber and Sandy, 1974).

Among the applications of Walsh functions are:

- Pattern recognition systems
- Image compression
- Design of radar waveforms
- Voice processing
- Signal multiplexing

Walsh analysis has several advantages over sinusoidal analysis.

- (1) Walsh functions contain real integral numbers of two levels ( $\pm 1$ ). This makes them directly compatible with digital integrated circuitry.
- (2) The Walsh function set is a natural describer of pulse-like signals. This property may be used to extract features from binary patterns (Alexandridis, 1971).
- (3) The transform from input data to Walsh function series expansion coefficients is faster than the FFT. For  $N$  samples, the FFT takes  $2N \log_2 N$  multiplications and additions (Blachman, 1974). The FWT takes only  $N \log_2 N$  additions (Harmuth, 1977).

Furthermore, techniques are available for obtaining two dimensional Walsh transforms in real time without waiting for all samples to be scanned before starting the transform operation (Alexandridis, 1971).

Three potential applications of Walsh functions are applicable in this context.

First, it is possible that the ETI signal may be a transmitted electromagnetic Walsh wave. The transmission of Walsh waves has been shown to have certain advantages in applications such as radar (Schreiber and Sandy, 1974). However, waves traveling through space would be filtered by the interstellar medium. A sinusoidal signal may be attenuated or phase-shifted, but the sinusoidal character would be preserved. A filtered Walsh wave would assume a significantly different characteristic.

Secondly, it is possible to analyze the MCSA output using a two-dimensional transform to pick out periodicities in time. The two-dimensional Walsh transform is attractive since the implementation of the fast Walsh transform is less expensive and faster than the FFT. However, for long samples of smooth signals, more terms are required in the Walsh series representation and greater accuracy is required of their coefficients for a given RMS total error (Blachman, 1974).

Finally, Walsh transform techniques are a possible approach to the recognition of binary patterns in the NBIT two-stage data compaction scheme. Axis symmetry properties of Walsh functions can be used in designing a feature extractor with a resulting reduction in dimensionality in the feature space (Alexandridis, 1971). However, due to the disadvantage of the Walsh transform being position dependent, position normalization of the binary pattern is necessary. In addition, specific

binary patterns which indicate a high probability of "signal presence" must be defined.

#### Conclusion:

Although the Walsh transform may lead to significant savings of hardware and processing time in analyzing binary or pulsed data, it seems not to be an all-purpose pattern recognizer. The Walsh transform was initially considered a candidate for serving as a "cartoon transform"; that is, a transform that would show strong narrowband response in the presence of an edge, outlined form, or filled-in pattern. Thus, while the axis symmetry properties of Walsh functions may be used to recognize specific patterns, but apparently provide no clue to the presence of signals of unknown characteristics.

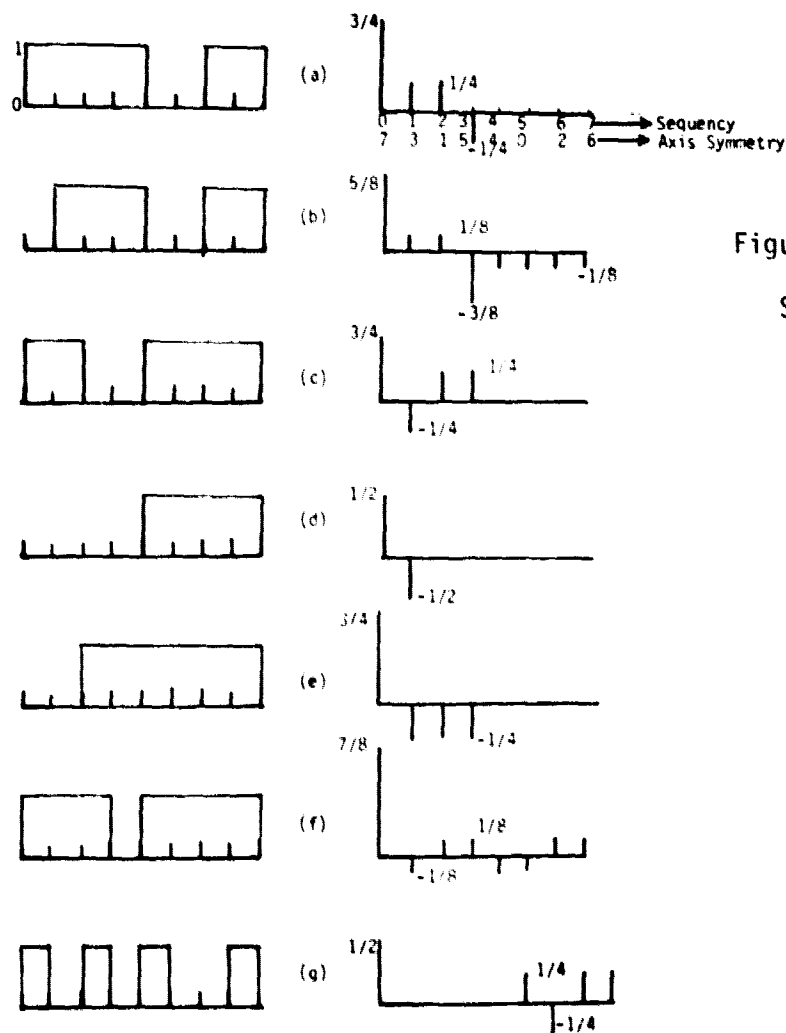


Figure A3.12.2 - Sequency Spectra of Various Patterns

## Appendix A4.1

### Large Memory Systems

#### A4.1.1 - Requirements

As is the case with many other dimensions of this study, the memory requirements of the proposed system are quite unusual. The two basic parameters of the memory requirements are the memory size and the memory read-write and access times that are required to handle the amount of data and the data rates that will be encountered.

One large memory system is required for the storage of the digitized input signal. The analog signals being received from each antenna feed are sampled at a  $16 \times 10^6$  Hz rate for the 1000-second observation period for each star, and quantized into 8-bit samples. This information must be stored, at least temporarily, in case a suspected ETI signal is discovered. The implication of this requirement is that storage (memory) is required for  $2.56 \times 10^{11}$  bits of digitized incoming data. These data arrive at a  $2.56 \times 10^8$  bits/second rate.

A second large memory system is necessarily involved with the output of the dual,  $8 \times 10^6$  channel multi-channel spectrum analyzer (MCSA). If all the data being output by the dual MCSA for 1000 seconds is to be stored temporarily for use by the ETI signal detection algorithms, the total required storage capacity is  $0.512 \times 10^{12}$  bits.

It rapidly becomes apparent that the hardware required to randomly access  $0.512 \times 10^{12}$  bits of memory in sufficient time to perform all the analyses used to detect ETI signals is prohibitively costly in dollars, power, and space requirements even if read-write and access time limitations allow realtime processing. For this reason, the output of the dual MCSA is truncated for some

of the algorithms to 4 bits each for the real and imaginary components of each feed; and further, we require temporary storage of "only" twenty seconds of MCSA output at a time. Thus, the data transfer rate into this memory is reduced to "only":

$$2 \times (4 + 4) \times 8 \times 10^6 = 1.28 \times 10^8 \text{ bits/second}$$

and a storage of "only":

$$20 \times 1.28 \times 10^8 = 2.56 \times 10^9 \text{ bits}$$

is required.

#### A4.1.2 - Types of Memories

The types of memories which are available to draw from for the various memory systems in our system are, at least theoretically, quite varied. These include:

- Semiconductor Random Access Memory (RAM)
- Magnetic Tape Systems
- Magnetic Disc Systems
- Read Only Memory (ROM)
- Charge Coupled Devices (CCD)
- Magnetic Bubble Memory (MBM)

Factors which eliminate some of these memory types from further consideration include the bit storage capacity of currently available units as well as those projected to be available by about 1982, current and future read-write rates, access times, and, of course, the present cost per bit of storage as well as projected future costs.

#### Less Conventional Memories

The last three types of memory listed above, the less conventional forms of memory, were eliminated from further consideration in this system for various reasons, including the desire to utilize a technology which has been



well-developed and proven to be a reliable means of data storage and retrieval.

Charge coupled devices (CCD) appear to be in the research and development stage continuously, and seldom get into production.

Magnetic bubble memories (MBM) seem to be moving ahead very rapidly with the immediate objective being to replace disc memories in the home computer market, (Mavity, 1979) and (Waller, 1979), in the very near future. However, from a cost-effectiveness viewpoint, bubbles were ruled out as a viable alternative at this time.

Read only memory (ROM) are employed in many of the ETI detection algorithms and/or as look-up tables, but they cannot play a role in the overall large memory system picture.

#### Exotic Memories

Additional exotic memory forms were investigated also. Optical recording has the capability of storing  $10^9$  bits per square inch. As attractive as this medium appears at first glance, it has the disadvantage of being a non-erasable medium, and thus it is usable only for archival storage in this project, (Hoagland, 1979), (Kenney, 1977), (Kaczorowski, 1977), (Kenville, 1978), and (Tufte, 1973).

Electron Beam Accessed Memory (EBAM) shows a good deal of potential, with densities of  $10^{12}$  bits per square inch possible, (Smith, 1978), (Speliotis, 1976), and (Hughes, 1975). The most serious fault with this technique is that to date no operating EBAM system has been built.

Holographic memories also present a possibility, but to date the memories are non-erasable and consequently are useful for archival storage only, (Gillis, 1975), (Harris, 1979) and (Honeywell, 1979).

### Random Access Memory

Recent activity involving solid state random access memory (RAM) has been primarily limited to decreased access time with an improvement by a factor of 5 having been achieved over the past two years, (Hewlett Packard, 1979). Another advance in RAM technology is the development of a very large semiconductor RAM-based memory by means of Adaptive Wafer Scale Integration (AWSI) as reported by Geideman (1978), Hsia (1979) and Brinton (1979). A  $1.2 \times 10^9$  bit version of the AWSI memory is in development and is scheduled for production within a year, with predicted cost to remain at about the present .03 cents/bit for semiconductor memory, (McDonnell-Douglas, 1979).

### Disc Memory

According to Hoagland (1979), the disc remains the primary form of mass storage. Despite competing technologies, improvements in disc technology will probably ensure its widespread use throughout the remainder of the century.

A disc system has been developed and is presently available which has multiple (20) read-write heads and recording surfaces which allow for 8 or 9 bits of parallel input or output data, (Ampex, 1979). This system can store  $2.4 \times 10^9$  bits and achieve  $8.7 \times 10^6$  bits/second data transfer rates.

As Hoagland (1979) also mentions, reduced cost per bit in all technologies derives primarily from an increase in density on the material being used for storage. If one were to double the bit packing density on a magnetic disc unit of the type described above, and then double the number of bits that are available in parallel for read-write operations, one would have the disc system presently being investigated in the Simulation Facility at Ampex, Redwood City, Calif. (1979). In any event, it is anticipated that the magnetic disc industry will be going to a double bit density magnetic medium. Also, the higher the

number of parallel outputs described above will be feasible in the near future. This system would have a capacity of  $4.8 \times 10^9$  bits per disc pack and a potential data transfer rate of  $3 \times 10^8$  bits/sec.

### Tape Memory

In terms of providing for the storage of massive amounts of data the most widely used medium is likely to continue to be magnetic tape. With increased recording densities and increased numbers of tracks per unit tape width, the amount of storage represented by a reel of magnetic tape is exceedingly large. The packing densities that are presently being utilized are on the order of  $10^6$  bits/square inch of tape (RCA, 1979), with one system under development which uses 180 tracks across a 2-inch magnetic tape. The result is a tape capacity on the order of  $5 \times 10^{11}$  bits, using a 2-inch, 10,000-foot tape (Bell and Howell, 1979).

A new tape system under development by Ampex (1979), which utilizes a flying head technology, is scheduled to be available by 1982. This system has a capacity of  $1.2 \times 10^{12}$  bits and is appropriately designated the Ampex Super HBR (high bit recorder).

In terms of rapid random access to stored data, the tape systems are useless, with access times to the center of the tape running up to two and a half minutes. But, in terms of quantity of storage, the tape unit is by far the least expensive storage medium.

Table A4.1 is a compilation of the characteristics of memory units available now and those envisioned to be available from three to six years hence.

	Character- istic	Mag. Tape	Mag. Disc	Solid State RAM		RAM-Based Memory System
				Static	Dynamic	
1979	Max Capac. per Unit (bits)	$4 \times 10^{11}$	$2.4 \times 10^9$	$1.6 \times 10^4$	$6.4 \times 10^4$	
	Read-Write Rate (bits/sec)	$6 \times 10^8$	$8.7 \times 10^7$	$1.6 \times 10^8$	$1.6 \times 10^8$	
	Access Time (sec)	* 120	** $2.8 \times 10^{-2}$	$10-50 \times 10^{-9}$	$100 \times 10^{-9}$	
	Cost (¢/bit)	$1.5 \times 10^{-4}$	0.003	0.1	0.03	
By 1982-1985	Max Capac. per unit (bits)	$1.2 \times 10^{12}$	$4.8 \times 10^9$	$6.4 \times 10^4$ (1985)	$1.28 \times 10^5$ (1983)	$1.2 \times 10^9$
	Read-Write Rate (bits/sec)	$7.5 \times 10^8$	$1.7-3 \times 10^8$	$1.6 \times 10^8$	$1.6 \times 10^8$	$1 \times 10^6$
	Access Time (sec)	* 150	** $2.8 \times 10^{-2}$	$10-50 \times 10^{-9}$	$80 \times 10^{-9}$	$1-2 \times 10^{-4}$
	Cost (¢/bit)	$2 \times 10^{-5}$	$1.5 \times 10^{-3}$	0.05	0.015	0.03

\* to middle of tape

\*\* average

Table A4.1 - Memory Characteristics

## References

\* PCP means Personal Conversation with Personnel

1. Ampex; PCP; Fall 1979, Topic: DM-PTD 9309 Disc System; Redwood City, CA.
2. Ampex; PCP; Fall 1979; Topic: Modifications to DM-PTD 9303 Disc System; Redwood City, CA.
3. Ampex; PCP; Fall 1979; Topic: Ampex Super HBR Tape System; Redwood City, CA.
4. Bell and Howell; PCP; Fall 1979; Topic: Bell and Howell System 300 and System 600 tape systems; Pasadena, CA.
5. Brinton, James B., Wafers to Challenge Disks, Bubbles?; Electronics, Aug. 16, 1979; 88-89.
6. Geidman, W.A. and A.L. Solomon; Wafer Integrated Semiconductor Mass Memory; 1978 International Telemetering Conference Paper.
7. Gillis, A.K. et al., Holographic Memories, Fantasy or Reality? IEEE National Computer Conference Paper; 1975.
8. Harris Corp; PCP; Fall 1979; Topic: Status of Optical Memories; Melbourne FL.
9. Hewlett Packard; PCP; Fall 1979; Topic: Hardware Availability, Present and Future; Palo Alto, CA.
10. Hoagland, Albert S., Storage Technology: Capabilities and Limitations; Computer, May, 1979.
11. Honeywell Corporation; PCP; Fall 1979; Topic: Status of Optical Memory Systems; Minneapolis, MN.
12. Hsia, Y. and G.C.C. Chang; Highly Reliable Semiconductor Mass Memory; 1979 NAECON Conference Paper.
13. Hughes, W.C., et al., BEAMOS, A New Electronic Digital Memory; 1975 National Computer Conference Paper.
14. Kaczorowski, E.M., Optical Mass Storage; 1977 IEEE COMPCON Paper.
15. Kenney, G. et al., An Optical Disc Data Recorder; 1977 IEEE COMPCON Paper.
16. Kenville, R.F., Optical Video Disc for Digital Mass Memory Applications; 1978 Spring COMPCON Paper.
17. Mavity, Willima C., Megabit Bubble Memories Move in on Mass Storage; Electronics, March 29, 1979; 99-103.

18. McDonnell-Douglas; PCP; Fall 1979; Topic: AWSI Memory System.
19. Smith, D.O., Electron Beam Accessed Memory; IEEE 1978 Spring COMPCON Paper.
20. Spiliotis, D.E., Electron Beam Addressable Memories; 13th Annual IEEE Computer Society International Conference Paper, 1976.
21. Tufte, O.N. and D. Chen, Optical Techniques for Data Storage; IEEE Spectrum, February, 1973.
22. Waller, Larry, Has Bubble Memory's Day Arrived?; Electronics, March 29, 1979; 80-81.

## Appendix 4.2 Discussion of Present and Alternative Architectures

In this appendix we will discuss briefly other signal detector architectures that might be more appropriate to MCSAs which output data at a rate  $\sim 10^2$  lower or  $\sim 10^2$  higher than the 8 Megawords per second rate of the Oasis System MCSA.

### A4.2.1. - Architectures for Smaller Scale MCSAs

With a throughput rate dropping from millions to thousands of data per second, many powerful processing possibilities are opened up. Floating point numbers rather than integers could be processed by programmable Floating Point Array Processors, now commercially available. The processors would be ideally suited to the high speed analysis of data because they possess the speed of dedicated hardware units and yet are programmable, and could be used to implement many of the algorithms described in this report. Their advantages are:

- Floating point preserves the dynamic range, eliminating clipping losses.
- Hardware building cost is traded for lower software designing cost.
- The software algorithms may be easily altered or replaced.

The reduced throughput of a smaller MCSA might also allow:

- Storage and random access of an entire observation on a few conventional disc drives (while  $\sim 20$  would be required for the 8 MHz MCSA).
- The possibility of human preprocessing, since a larger percent of an observation could be viewed with a display.
- The data of an observation could be stored in RAM memories (say 64 Kbits each) and a network of minicomputers could simultaneously operate on each RAM employing either identical or varied detection algorithms, i.e. network processing.

### A4.2.2 - Architecture for Larger Scale MCSAs

For MCSAs on the order of 300 Mega-Channels per second, processing of the entire array with the algorithms of the Oasis system becomes nearly impossible. A different approach would be to sample and analyze the output of selected

"groups" of channels sequentially. In other words, if we limit the total search to one for persistent carrier waves, pulses, etc., then we need only check a set of adjacent channels periodically for the presence of such a signal. The processing required can thus be arbitrarily reduced, but the price for this reduction is a loss of sensitivity to ephemeral signals. Thus, for the Oasis system to analyze the data from a spectrometer with  $300 \cdot 10^6$  channels output per second at an 8 MHz rate, it is first necessary to consider every 40th block in the first row of blocks, and then to consider each 40th block offset by one in the second row of blocks. Random selection of blocks may be more desirable so as to avoid systematic omission of identically drifting spectral features.

Finally, the signal detection algorithms for such a large bandwidth spectrometer might be justifiably limited to simpler tests, such as employed by the pulse detection algorithm, and the total power detection algorithm of the NBIT processor.

#### A4.2.3 - Critique of the Oasis Design: Weaknesses

The proposed Oasis System is a rather complex system. The Summer Study Group as a whole consciously declined from rejecting any potentially powerful algorithm simply on the basis that it made for a more complex final system. Every attempt was made to evaluate the relative merits of each algorithm, yet time did not allow for a truly fair and complete evaluation and ordering of the worth and merits (both as signal-specific, and signal-arbitrary tests) of the dozen or so different algorithms implemented. While much work was done toward this end, much more remains to be done. It is recommended that further computer simulations of these algorithms be conducted, upon a more varied (in SNR as well as type) set of earth-civilization-type signals as well as artificially concocted signals in noise. If nothing else, such a study would yield the relative weights to place on the alarm thresholds, as discussed in Chapter 5. To the hardware designer wishing to construct a sensitive, yet simpler and cheaper



system, the Oasis System, as it stands, represents a menu or shopping list of algorithms. This hypothetical designer is in turn given from this report the plans necessary for implementing and integrating any subset of the algorithms deemed most effective.

The decision to retain 16 output levels (4 bits) of real and 16 levels of imaginary signal amplitude, for each polarization in the Oasis System was made without a thorough evaluation of the resulting clipping losses inflicted on each algorithm. The decision was based instead on the high cost of providing another set of 4 high density disc drives for, say, 8 bits of amplitude resolution. The GCV and ANOVA algorithms were tested through simulations for clipping losses and were found to be only slightly degraded by 4-bit clipping (see the appendices). Similarly, the proposed simulation of all Oasis algorithms should include a check for the clipping sensitivity of each.

#### A4.2.4 - Critique of the Oasis Design: Strengths

Chapter 3 has described the first and second stage processing of the entire  $8 \times 10^6 \times 10^3$  element array. It detailed how the array was divided into blocks, each tested and compacted into 9-bit words forming the elements of pattern blocks for pattern recognition and cluster detection. The scale of the array subdivision test is gradually increased in each of these stages. This is shown in Table A4.2.1.

Table A.4.2.1 The Increasing Scope of the Major Search Algorithms

Algorithm	Items Tested	Number of Array Elements Considered at Once
Pulse Detector (1st pass)	individual elements	$10^0$
Carrier Wave Detector	60 elements/slice*	$10^{1.7}$
NBIT Tests	20 x 40 element block	$10^{2.9}$
Carrier Wave Detector (Accumulated)	3000 element slice	$10^{3.5}$
Cluster Detector	200 x 50 elements, each a block	$10^{6.9}$
(Size of Entire Array for comparison)	$8 \times 10^6 \times 1000$	$10^{9.9}$

\* each slice includes 3 elements per row x 20 rows of the entire array

This slowly increasing scale or scope of the Oasis search, especially with such signal-arbitrary algorithms as NBIT and Cluster detection is considered a major strength of the system. This table also implicitly gives a rationale for the sizes chosen for Block and Pattern Block arrays.

A second strength of the system is found in the transpose hardware system which, in real time, arranges the data into a much more desired form within a memory unit. To search effectively for any signal, an individual channel's output over time must be stored in a memory device contiguously and analyzed in a like manner. The array transposer accomplishes this for the algorithms, for the archives, as well as for the operator's perusal with an advanced image processor.

Another strength of the system is felt to be found in the Generalized Coherence measure. This algorithm or measure seems to implicitly embody many fundamental SETI principles while providing a new means of compressing the output of an MCSA. Since it is a succinct measure of polarization, phase, and amplitude coherence of received time series, it serves as a detector for those qualities expected to be characteristic of intelligent communication.

Finally, the Oasis System places a good deal of emphasis on the capability of the operator to detect and identify signals. For this reason, the role of the operator has been given a wide latitude, including his/her facility of viewing and interacting with all stages of the hardware processing. Indeed, to use the operator as the learning part of the system, capable of modifying the performance of the rest of the system and possessing complete control over the all-important archiving decision, we believe greatly enhances the performance of the total system.

### Appendix A4.3

#### Implementation for (1) Degree of Polarization and

#### (2) Broadband Pulse Detection (with GCV)

Since these two algorithms require some common computation, they share a common front end processor which passes the values on to the two subroutines: m-routine, and GCV-routine. Conceptually, there are three processing components: the front end processor receives the incoming data, computes and accumulates the power associated with each polarization, the cross product of the two polarizations, for corresponding data points, and the auto- and cross products for two adjacent frequency data points. After 40 frequencies have been processed, the first three sums are sent to the m-routine which calculates the degree of polarization. The other sums are passed along to the GCV routine. In addition, the front end processor computes three other factors used by the GCV routine: normalized power factors for each polarization and a phase coherence angle ( $\delta$ ).

The m-routine calculates the degree of polarization in each 40 Hz row of data according to the formula:

$$m = \frac{\sqrt{(\sum |x_i|^2 - \sum |y_i|^2)^2 + 4 |\sum x_i y_i^*|^2}}{(\sum |x_i|^2 + \sum |y_i|^2)}$$

The result is compared to a predetermined threshold and, if it exceeds the threshold, the routine sets on the 1-bit and sets an alarm for the central processor.

The GCV routine calculates the general coherence value for each 40 Hz row of data according to the formula:

$$GCV = \frac{1}{N-1} \left| ( K_1^2 \sum x_{i+1} x_i^* + K_1 K_2 (e^{j\delta} \sum x_{i+1} y_i^* + e^{-j\delta} \sum y_{i+1} x_i^*) + K_2^2 \sum y_{i+1} y_i^* ) \right|$$

where

$$K_1 = \frac{1}{N} \sum \frac{|x_i|}{\sqrt{|x_i|^2 + |y_i|^2}} \quad K_2 = \frac{1}{N} \sum \frac{|y_i|}{\sqrt{|x_i|^2 + |y_i|^2}}$$

and

$$\delta = \tan^{-1} \frac{\text{Im } \sum x_i y_i^*}{\text{Re } \sum x_i y_i^*}$$

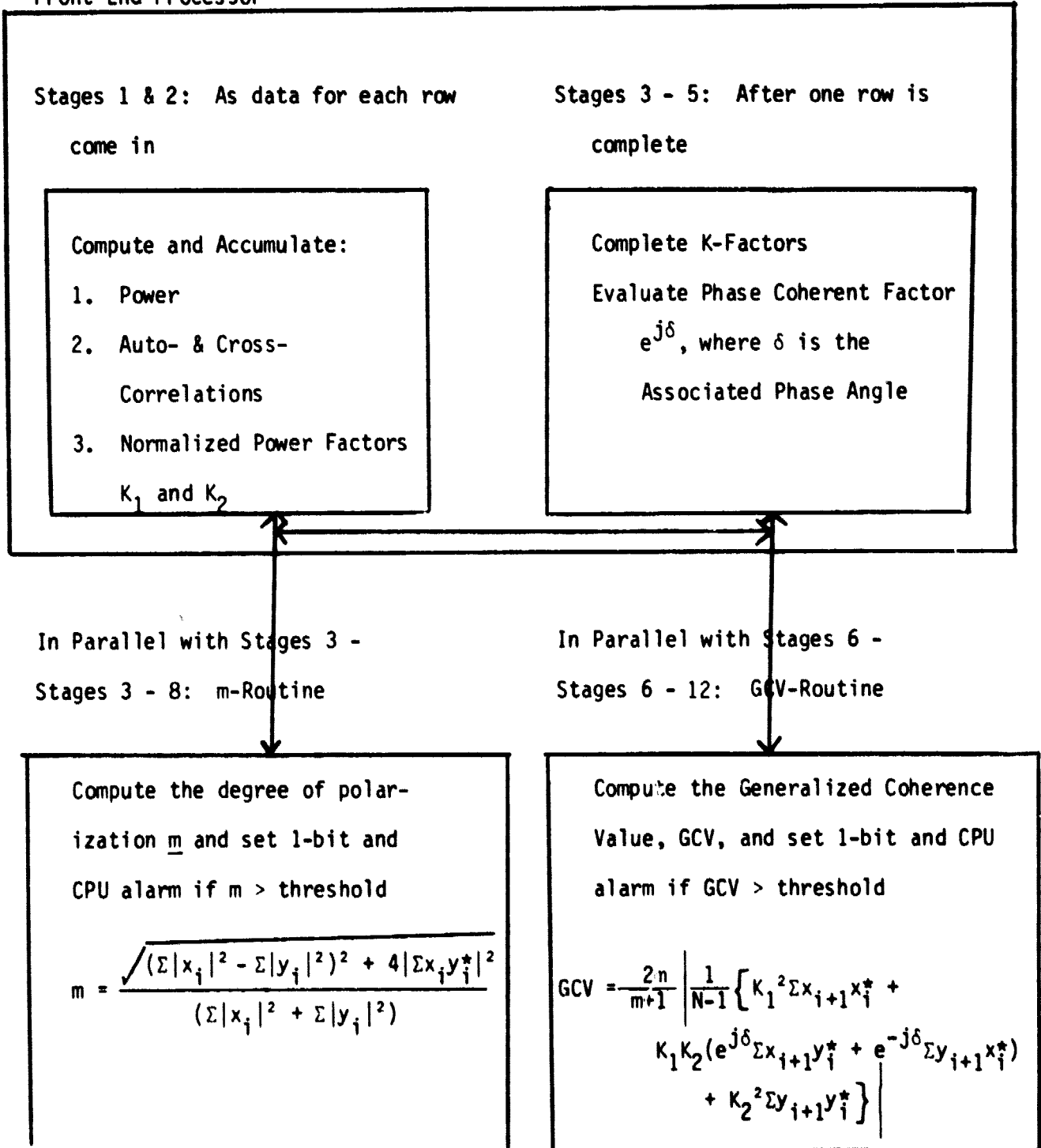
Part of the processor is operating on the incoming data while the rest of the calculation is being processed in parallel. The entire processor is a multi-stage pipeline. It is outlined in Figure A4.3.1 and presented in Figures A4.3.2 - A4.3.4.

Figure A4.3.1

Schematic Block Diagram for:

- (1) Degree of Polarization by Row and
- (2) Degree of Generalized Coherence by Row

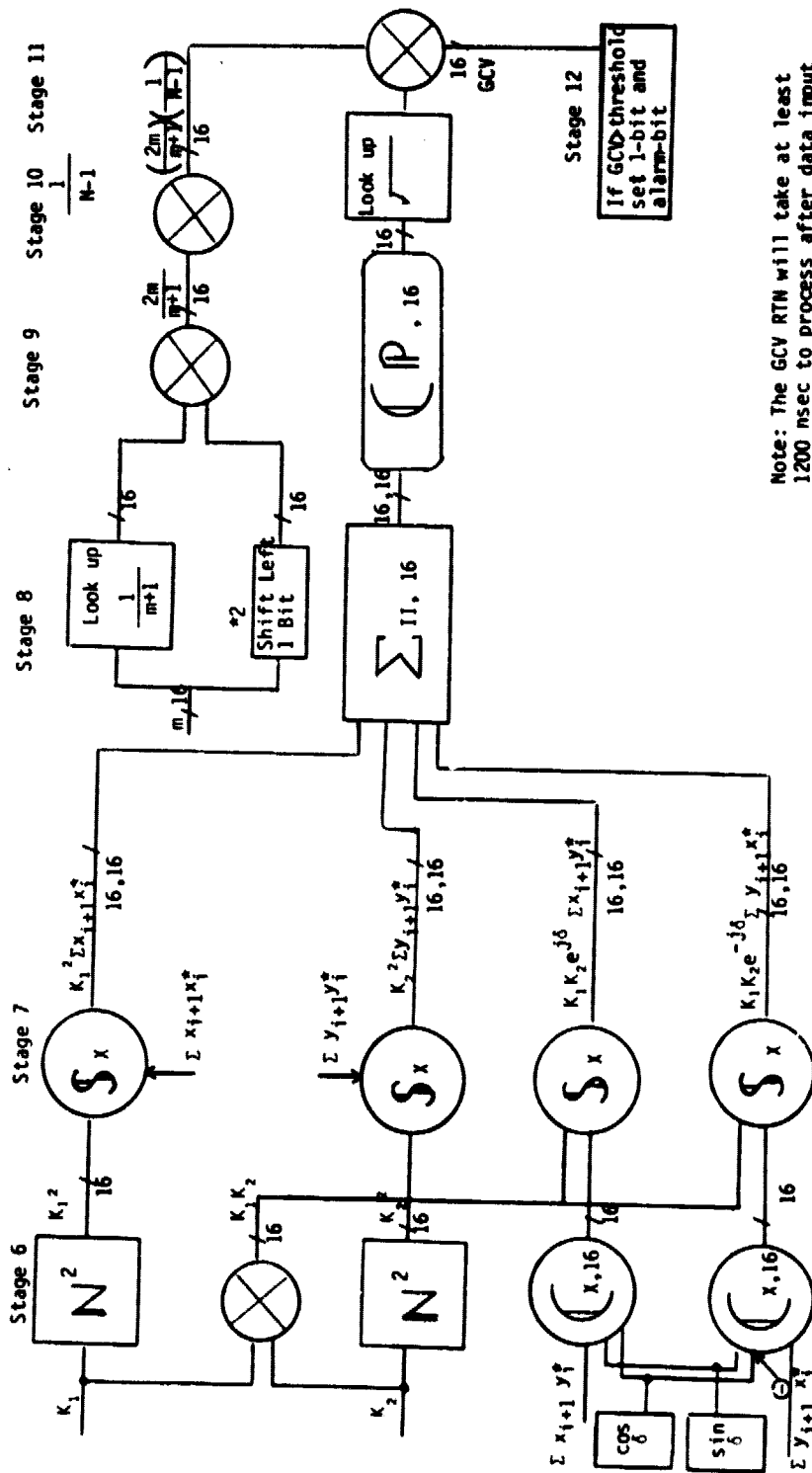
Front End Processor





$$\text{GCV - Routine GCV} = \frac{2m}{m+1} \left| \frac{1}{N-1} \left\{ K_1^2 \sum x_{i+1} x_i^* + K_1 K_2 e^{j\delta} \sum x_{i+1} y_i^* + K_1 K_2 e^{-j\delta} \sum y_{i+1} x_i^* + K_2^2 \sum y_{i+1} y_i^* \right\} \right|$$

$$K_1 = \frac{1}{N} \frac{\sum |x_i|}{\sum |x_i| + \sum |y_i|} \quad K_2 = \frac{1}{N} \frac{\sum |y_i|}{\sum |x_i| + \sum |y_i|} \quad \delta = \tan^{-1} \frac{\text{Im} \sum x_i y_i^*}{\text{Re} \sum x_i y_i^*}$$



Note: The GCV RTN will take at least 1200 msec to process after data input is complete.

Figure A4.3.3 - Second Stage of GCV by Rows

m-Routine

$$m = \frac{\sqrt{(\sum |x_i|^2 - \sum |y_i|^2)^2 + 4 |\sum x_i y_i|^2}}{(\sum |x_i|^2 + \sum |y_i|^2)}$$

Degree of Polarization

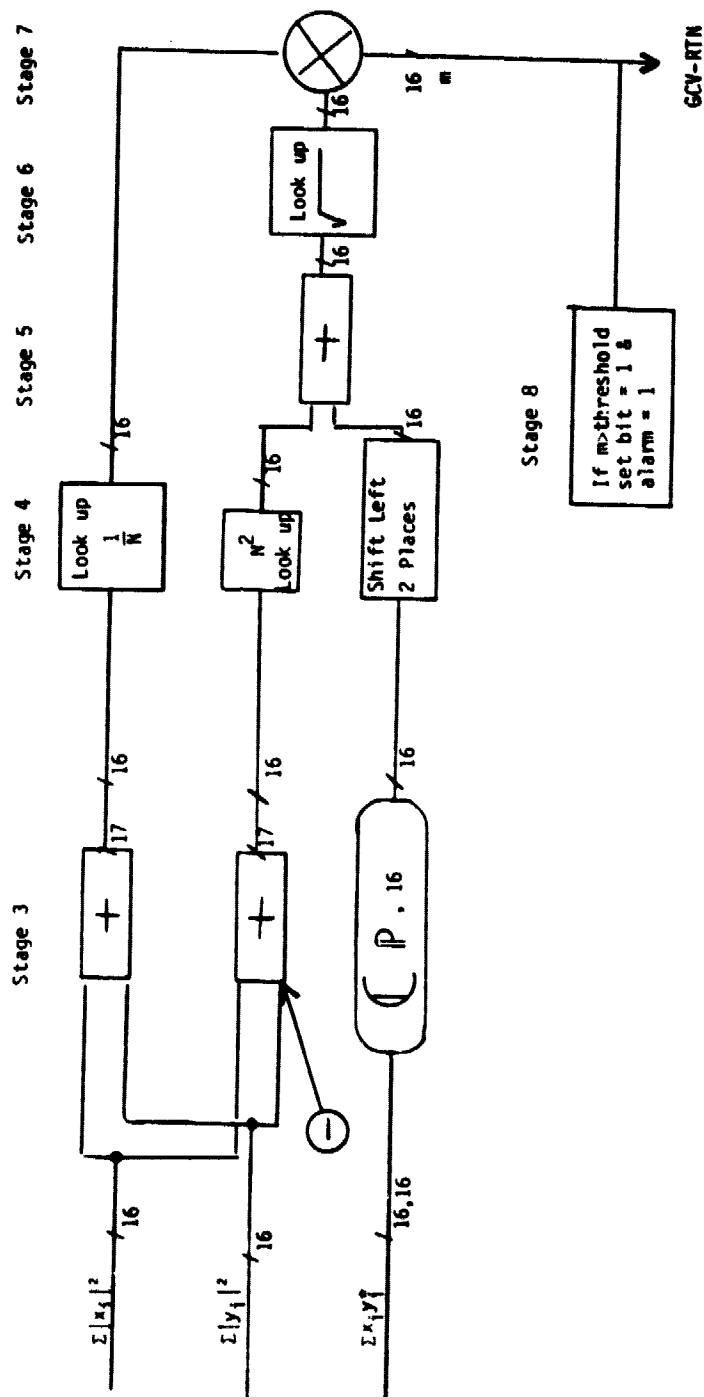


Figure A4.3.4 - Parallel GCV by Row  
Second Stage also computing MBIT Degree of Polarization



#### Appendix A4.4

##### Survey of Image Processors

This appendix summarizes the results of a survey of commercially available digital image processors. The evaluation of each system was based upon its potential role as the operator-machine interface in the Oasis signal detector. The survey included careful examination of manufacturers' literature, personal conversations, and demonstration of operational image processing systems. Two previous surveys exist. (LaPado, R., Reader, C., Hubble, L., ESL engineering report, and Allen, R.J., and ter Horst, H.J., Excerpt from Market Study of Several Image Processing and Display Systems, University of Groningen).

The current trend in state-of-the-art image processor architecture is a bus-oriented, pipeline array processor utilizing function tables and memory multiplexing techniques to maintain video speed (10 MHz) while performing various operations on image data.

The most powerful image processors are capable of stand-alone operation with a keyboard and mass storage devices controlled by an on-board micro-processor. More commonly, the image processor will be interfaced to a host mini-computer for additional computational and I/O capabilities.

##### Refresh Memory

The image data resides in refresh memory that is typically dual-ported RAM arranged in arrays of 8-bit pixels for 256 intensity levels per pixel. A monochrome CRT requires one refresh array, and a color CRT is refreshed from 3 arrays, one each for red, green, and blue. The size of the image data base is limited only by the address space. This varies among manufacturers

and is usually more than can be displayed at once because of amplifier bandwidth limitations in the CRT. Scrolling overcomes this limitation by allowing the user to roam at will through the entire refresh memory.

In addition to image refresh memory there is usually one or more channels of graphic overlay memory. Histograms, vectors and alphanumerics are stored in this memory and combined at the processor output by pixel replacement or color combination so they are visible against the image background.

### Pipeline Processors

Refresh memory contents are read out and processed before conversion to analog video signals. The use of function table look-ups and high speed arithmetic hardware is necessary to maintain video refresh rates. Pipelined processing allows an operator to scroll, zoom on a specified area without altering refresh memory contents, obtain pixel statistics, and perform image combination arithmetic in one or a few refresh times. A feedback loop allows processed output to be written back into refresh memory so that more complex iterative functions such as spatial convolution and filtering can be accomplished in a few seconds.

### Displays

The display device for high resolution monochrome or color monitors uses refresh rates of 30 Hz interlaced or 60 Hz non-interlaced. Some image processors accept external sync for use with video discs or digitized TV camera input. The highest spatial resolution in both monochrome and color monitors

currently available is 1024 x 1024 pixels. Some of the most recently introduced image processors will handle the higher data rates for these monitors as well as standard 512 x 512 pixel monitors. Three image processors--the Comtal Vision one/20, the De Anza IP 5000 and the I<sup>2</sup>S Model 70--meet the requirement of the Oasis system. A comparison of the more important features of these image processors is summarized in Table 4.4.1, and in the following paragraphs.

#### Comtal Vision one/20:

- Refresh memory can be expanded to 134 Mbits. Memory configuration is dynamic and under firmware control.
- Can accomodate up to 4 independent user stations.
- Multiple images in refresh memory can be displayed in rapid succession resulting in a "loop movie" effect.

#### De Anza IP 5000:

- Good processing and feedback loop.
- 3 x 3 convolution in less than one second.
- Feedback is routable to any bits of any refresh memory by defining a 32-bit mask.

#### I<sup>2</sup>S Model 70:

- Extensive software is available, such as utilities for file management and image processor control and a library of image enhancement and transformation routines.
- A feedback ALU does 16-bit arithmetic on images.
- Minimum and maximum pixel values and other pixel statistics are continuously available.

	De Anza IP 5000	Comtal Vision One/20	I <sup>2</sup> S Model 70
Software	diagnostics, limited utilities, DEC compatible. Mostly Fortran	none	extensive image processing and utility library. DEC compatible, mostly Fortran.
Zoom	non-destructive pixel replication. x2,4,8 magnification	non-destructive bi-lateral interpolation. x2,4 magnification	non-destructive pixel replication. x2,4,8 magnification
Scroll	up-down, left-right in 1-pixel increments	up-down, left-right in 1-pixel increments	up-down, left-right in 2-pixel increments
Split-Screen	rectangular portions of 2 or 4 images	regular or irregular portions of images superimposed on display	rectangular portions of 2 or 4-pixel increments
Refresh Memory	one to four 640 x 512 8-bit images. Readback to host.	800 ns RAM expandable to 64 512 x 512 8-bit images. Readback to host.	I/O up to 2M pixels/sec 1 to 12 512 x 512 8-bit images addressable by row or column. Readback to host.
Graphics Overlays	8 512 x 512 x 1 bit overlays. Scrollable.	any part of refresh memory dynamically assignable	8 512 x 512 1-bit overlays. Scrollable.
Inter-active Controls	trackball, tablet, function buttons, joystick	trackball, tablet, function buttons	trackball, tablet, function buttons, joystick
Cursor	16 x 16 pixel programmable	15 x 15 pixel programmable	64 x 64 pixel programmable

- continued -

	De Anza IP 5000	Comtal Vision One/20	I <sup>2</sup> S Model 70
Function Tables	1 per image	8 per image	3 per image, 1 input table
Micro- Processor	LSI-11	LSI-11	LSI-11

Table 4.4.1 - Image Processor Comparison

### References

1. Allen, R.J., and Ter Horst, H.J., Excerpt from a Market Study of Several Image Processing and Display Systems, unpublished. Astronomy Dept. and Computer Center, University of Groningen, The Netherlands, Dec. 1977.
2. LaPado, R., Reader, C., and Hubble, L., Image Processing Displays: A Report on Commercially Available State-of-the-Art Features, unpublished ESL engineering report no. ER-219, Aug. 21, 1978, Sunnyvale, CA.

## Appendix A5.1

### Experiments on Observer Detection and Recognition of Simulated Signals in Noise

Perhaps the most critical information we desire concerning the human observers is an understanding of his/her ability to detect weak signals amidst background noise. There is an extensive literature in psychology on this problem, but most of this work deals with the case where the signal characteristics are known by both the subject and the experiment (see Swets, Green, Getty, and Swets, 1978 for examples in the visual modality). The challenging aspect of the SETI problem is that nobody has a firm idea concerning the precise nature of the signal. Moreover, it appears that little research has been done using visual noise fields similar to the ones anticipated in the present context (except Swets et al., 1978). For these reasons, then, we conducted some preliminary experiments with simulated visual displays. Observers made estimates of their confidence that a signal was present under a variety of conditions, spanning several signal types at various signal to noise ratios, as well as complete noise fields. Further, we present data to a new technique whereby observers draw potential signals on a matrix representing the visual display. Analysis of such drawings may permit us to infer the accuracy of localizing detections, and the types of "pseudo" signals observers perceive in noise fields.

Observers. The participants were 14 members of the summer study program and 4 other professionals at the NASA-Ames research laboratories. They each participated on a voluntary basis, and were tested individually.

Equipment. The display simulations were generated by a Hewlett-Packard 9800 computing system consisting of two primary components: the 9825A

calculator and a black and white video display unit. The video display was built especially by Hewlett Packard for the SETI program. It presents a 128 x 128 pixel noise field with 256 (8 bit) levels of intensity. The entire screen must be erased before the next picture is written; refresh time for the screen is 30 frames per second.

Displays. The noise was generated over  $2^7$  levels of amplitude approximating a Rayleigh distribution. These amplitude values then were squared to approximate an exponential distribution of pixel intensities on the screen. This noise field is what the observer saw in all displays (different noise fields were used for each presentation).

Signals were superimposed on the noise fields giving the impression of a pattern of brighter pixels. Average signal to noise ratios (SNR) were .4, .6, .8, 1, 1.2, and 1.4.

Experimental design. Three patterns were employed at each SNR: straight line, wavy line, and pulse. The shape and calculation of these patterns are presented in Figure A5.1.1 (panels a, b, c). Examples of each pattern at the highest SNR, as seen on the screen, are shown in Figure A5.1.2.

The starting point of the signal was randomly selected from positions across the top edge of the screen and extended to the bottom edge. In addition to the signal displays (6 presentations for each of 3 signal conditions, for a total of 18) six noise field displays (i.e., without signals) were used. Hence, there were 24 displays presented to each observer.



Procedure. The 24 displays first were presented in a random order for approximately 10 seconds each. This phase was introduced to give the observer an idea about the range and type of displays. Next, the series was repeated and the observer gave a confidence rating from 1 to 10 concerning the presence of a signal anywhere in the display. Part of the instructions for this phase read: "We want you to estimate your confidence that a signal (of any shape, size, or type) is present. Give a number from 1 to 10 where 1 represents minimum confidence and 10 represents maximum confidence." The procedure was self-paced; the display was changed automatically after the observer's estimate was recorded by the experimenter.

In the third phase of the experiment, the series was repeated and confidence ratings obtained. In addition, observers drew potential signals on a sheet of graph paper. This page was divided into 64 equal "boxes" or zones (i.e., 8 x 8 matrix). The instructions for this phase read (in part:

"In this last phase first give a confidence rating for each display. Don't try to recall what number you gave last time. Make each judgment independent of earlier ones. After giving a confidence rating for a display, draw potential signals on this graph paper. Precision is not important, just try to localize the perceived signal as best you can. For each display you must draw at least one signal regardless of your confidence rating. In addition, you may draw other signals for the same display. These can be short lines, pulses, or other shapes, located anywhere in the display."

The experiment was conducted under low room illumination which produced satisfactory visual contrast levels for the displays. The entire set of observations for a single session took approximately 40 minutes.

## Results and discussion

Preliminary analysis suggested that there was minimum difference between confidence ratings on the two trials (phases), so these data were combined in further analyses. The means and standard deviations of the confidence ratings were calculated for each of the 24 displays. The results are presented in Table A5.1.1 (as a function of SNR and pattern type: line, wavy, pulse). The mean data are presented graphically in Figure A5.1.2.

Turning first to the mean ratings, there are obvious differences in an observer's confidence of his ability to detect the three types of patterns. The average rating for the noise fields is approximately 2.5 on the 10 point scale. Using this value as a lower referent, linear patterns are first detected at a SNR of .8; pulses and wavy lines at a higher value (1.0). Hence, by this measure, an observer's absolute detection threshold is lowest for straight lines. Further, once a straight line is detected, its presence is obvious, as indicated by the sharp increase in confidence at  $\text{SNR} = .8$ . On the other hand, the confidence ratings for the wavy lines and pulses increase gradually, and level off at a low value than for straight lines. In short, straight lines are the easiest to detect; pulses and wavy lines are about equal in detectability across the values of SNR tested.

The standard deviations also depend on the signal pattern. In general, the standard deviations (see Table A5.1.1) are larger for the pulses and wavy lines than for the straight lines. The standard deviations also vary with SNR; greater variability was obtained for the higher SNR values with the exception of the straight lines, where variability drops dramatically at large SNR values.

The finding of an increase in variability with higher mean ratings at first glance appears paradoxical. However, examination of the data of individual observers explains this contradiction. What appears to have happened here was that observers split into two distinct groups: one group definitely saw the signal, the other appeared to miss it completely. Hence, the combined data are highly variable.

The signal drawing data were analyzed by calculating the percentage of hits and false alarms for each stimulus pattern. A hit is defined as any mark drawn in a cell of the 8 x 8 matrix that in fact contained part of a signal. The percentage of hits is the total number of correctly marked cells for all 18 observers divided by the total signal cells (multiplied by the number of observers). Since this measure by itself may or may not reflect accuracy (for example, a perfect score would result if observers marked every cell of the matrix), it is necessary to compute the false alarms as well. The percentage of false alarms is defined as the total number of marked cells, where a signal was not present, divided by the total number of non-signal cells in the matrix (multiplied appropriately to obtain a single group score).

The percentage of hits is shown in Figure A5.1.3 as a function of signal to noise ratio for each of the three signal types. The trend of the results is very similar to the confidence ratings presented in Figure A5.1.2. The hits for the linear signal show a dramatic increase at SNR = .8; whereas the trend is more gradual (and in fact, non-monotonic) for wavy lines and pulses.

The false alarm results are shown in Figure A5.1.3. For each signal pattern the false alarm rate is at or below the rate for noise alone (~ 15-20%).

The linear signal yields the lowest false alarm rate. Quite obviously, the hit results are not artifactually due to a high percentage of the matrix being marked independent of signal presence.

Histograms were constructed to represent the location of drawn signals aggregated for all observers. Figure A5.1.7 shows the frequency of drawn marks for the linear pattern at the highest SNR (1.4). It is clear that observers show substantial agreement on the location of the signal. (For purposes of representing the data, the size of the matrix on the histogram is larger than the matrix used in the experiment.)

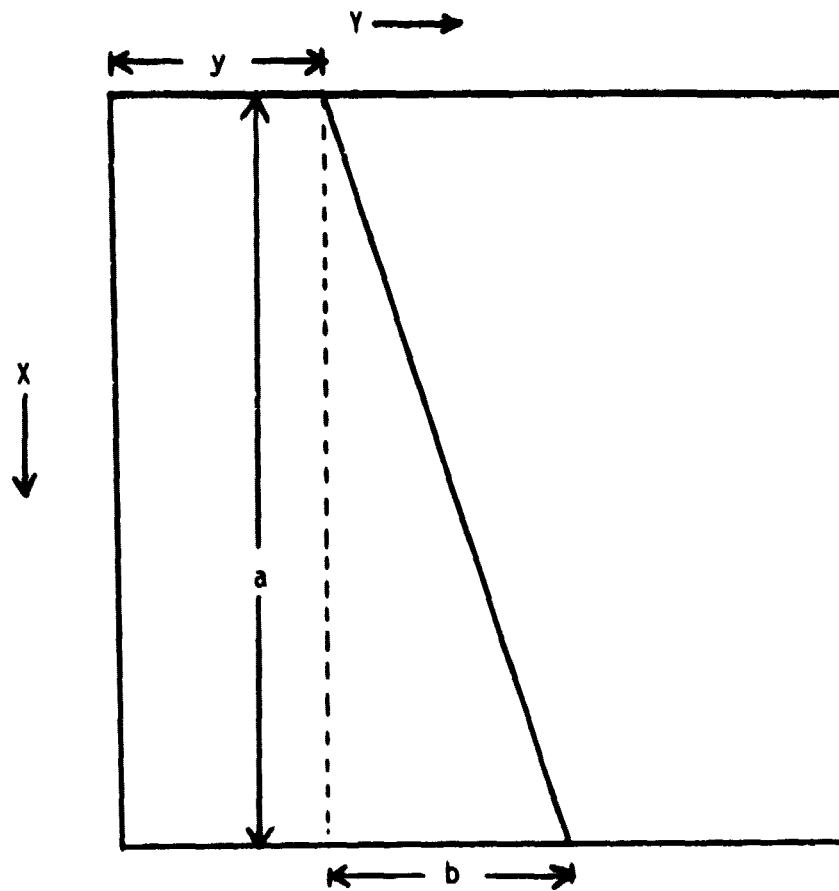
Figures A5.1.8 and 9 present results for two noise fields. The interesting aspect of these histograms is the apparent differences. In Fig. A5.1.8 the agreement is minimal among observers concerning the location of potential signals. However, a fair amount of agreement is evident in Figure A5.1.9. The latter finding is intriguing since it suggests that characteristics of a noise display lead people to infer the presence of signals, even though their confidence in signal presence is rather low (the mean confidence ratings for the two patterns are approximately the same).

The findings here must be considered tentative. Further analysis of individual differences among observers, with replications using larger sample sizes, should help to clarify and extend the present study along avenues that invite direct application to the problem of signal detection and recognition.

S N R		0	.4	.6	.8	1.0	1.2	1.4
Linear	$\bar{X}$	2.44	2.08	2.25	8.61	8.0	9.31	9.69
	$\bar{S}$	1.88	1.32	1.44	1.79	2.18	1.34	.82
Pulse	$\bar{X}$	—	3.08	2.47	3.03	3.64	6.31	6.58
	S	—	2.52	1.96	2.31	2.47	2.96	2.88
Wavy	$\bar{X}$	—	2.50	2.69	2.61	3.44	5.14	7.33
	S	—	1.90	2.05	2.25	2.27	3.25	2.72

Table A5.1.1 - Composite Statistics for Confidence Ratings

Linear Signal



$y$  = Frequency Offset

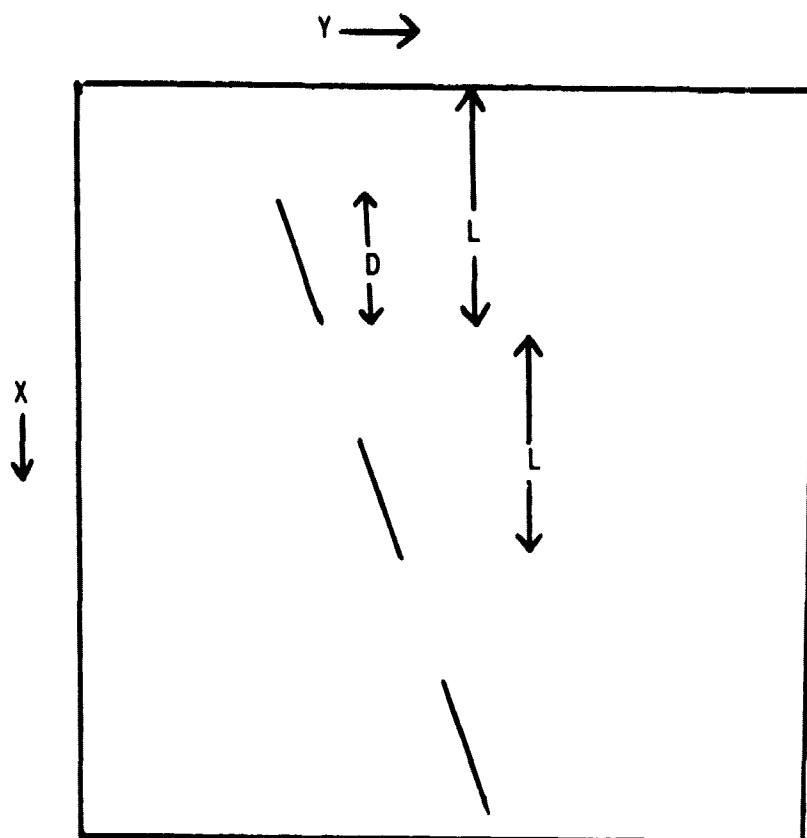
$M = \frac{b}{a}$  = Slope

Function:

$$Y = MX + y$$

Figure A5.1.1 (a)

# Pulsed Signal



$$Y = MX + y \quad (\text{see Linear Signal})$$

$$\text{If } X > (D + L \text{ int}(X/L)) \quad \text{then } Z = 0$$

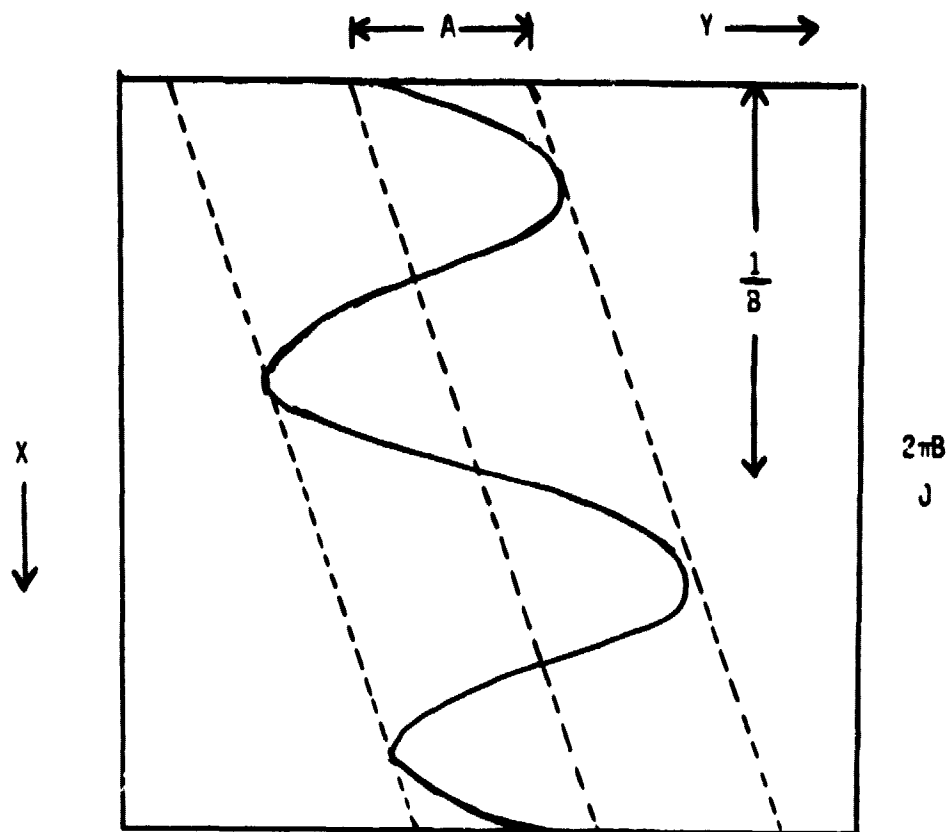
$Z$  = Intensity

$L$  = Pulse Length

$D$  = Pulse Duration

Figure A5.1.1 (b)

# Wavy Line



$$Y = \underbrace{MX + y}_{\text{see Linear Case}} + A \sin(2\pi BX)$$

$A$  = Modulating Limit

$B$  = Modulation Frequency

Figure A5.1.1 (c)



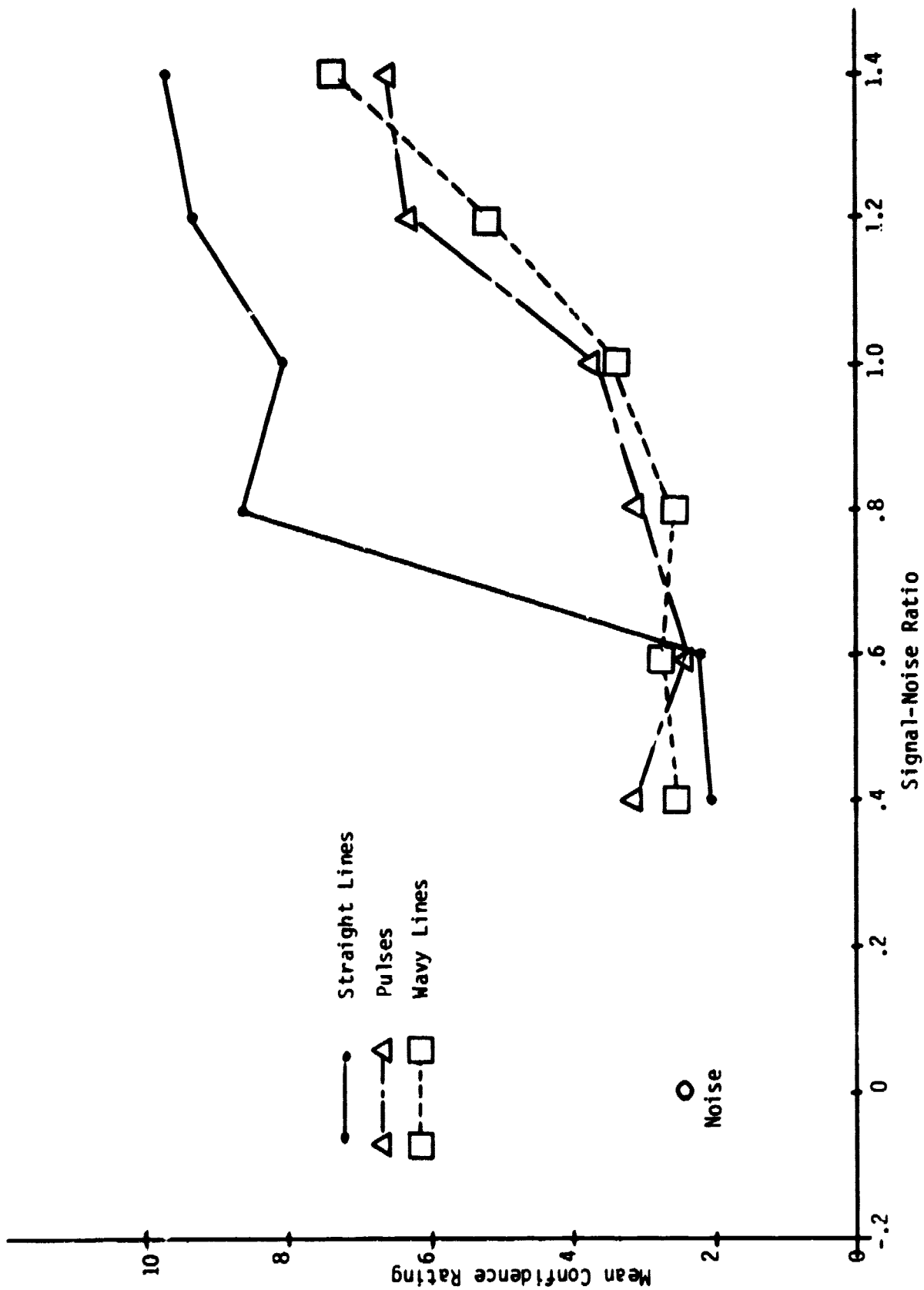


Figure A5.1.2

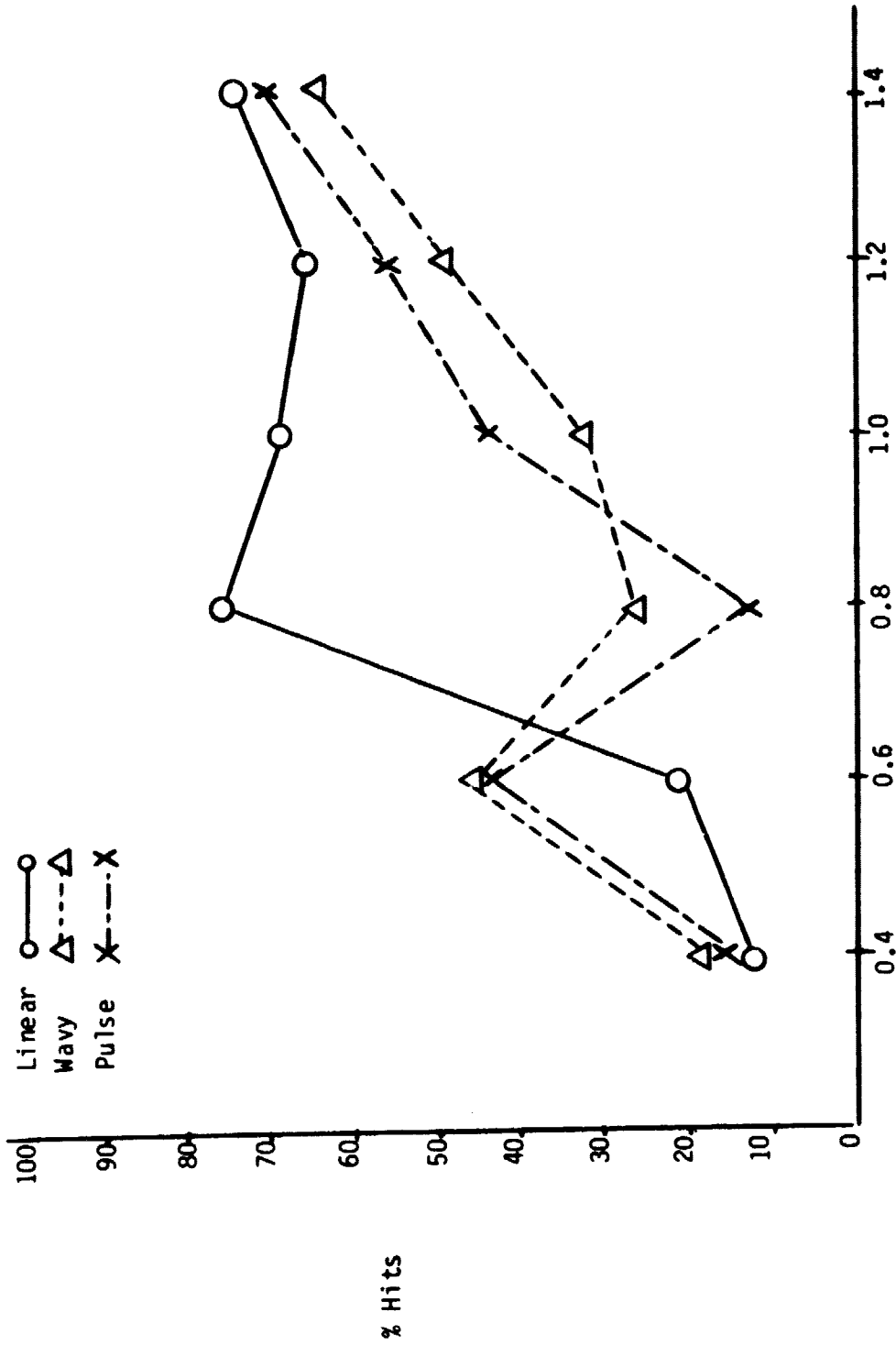


Figure A5.1.3

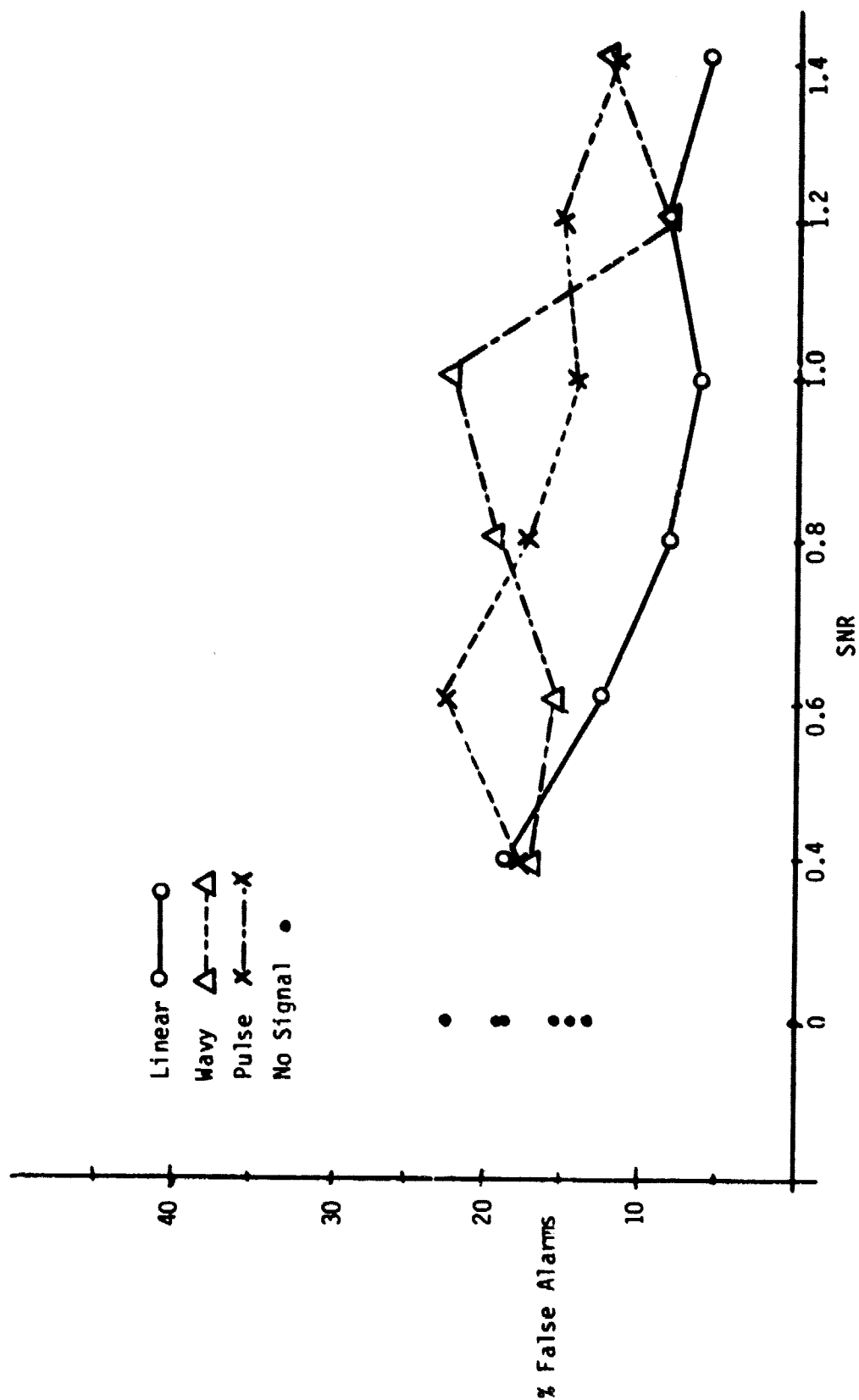


Figure A5.1.4

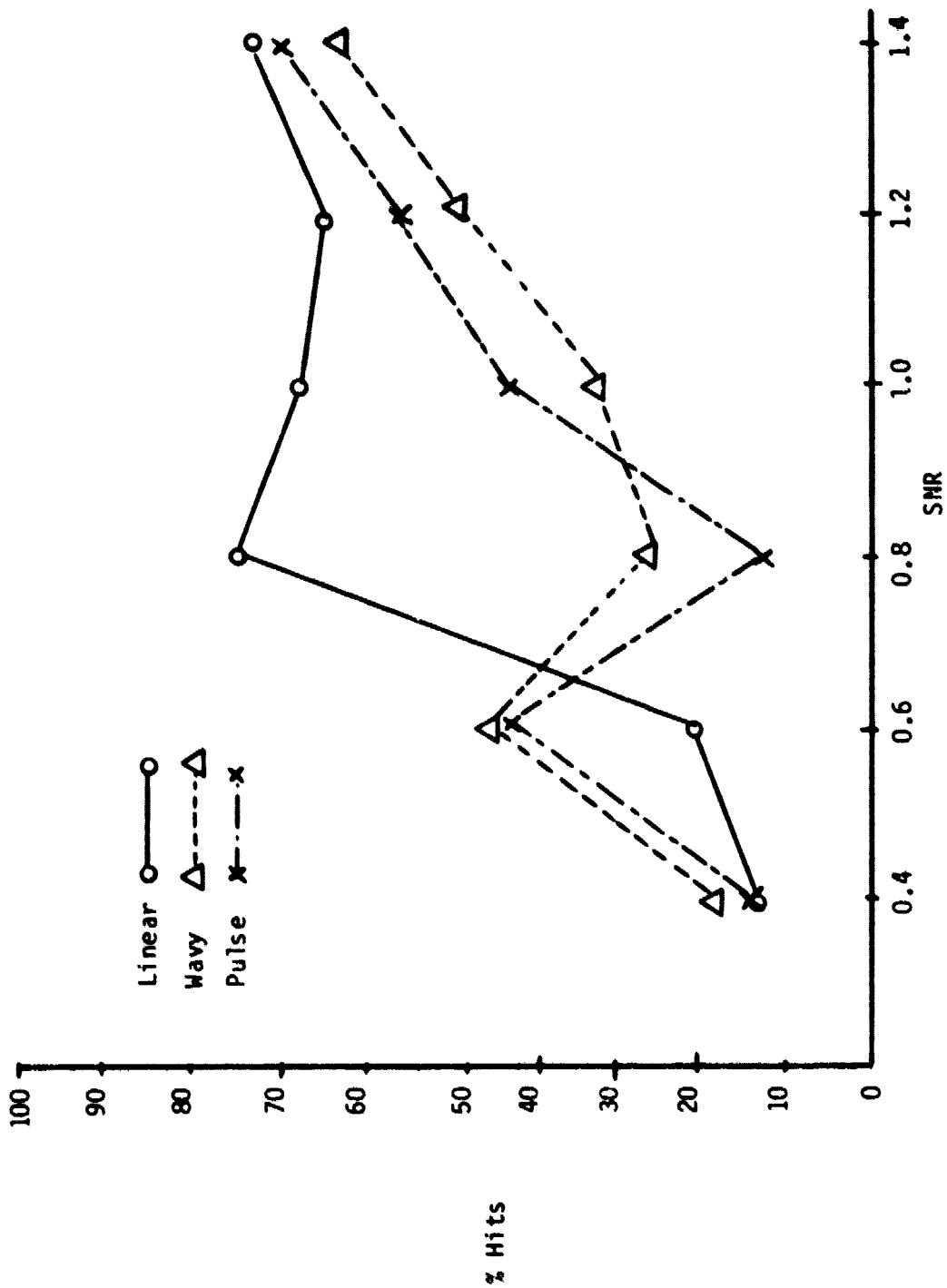


Figure A5.1.5

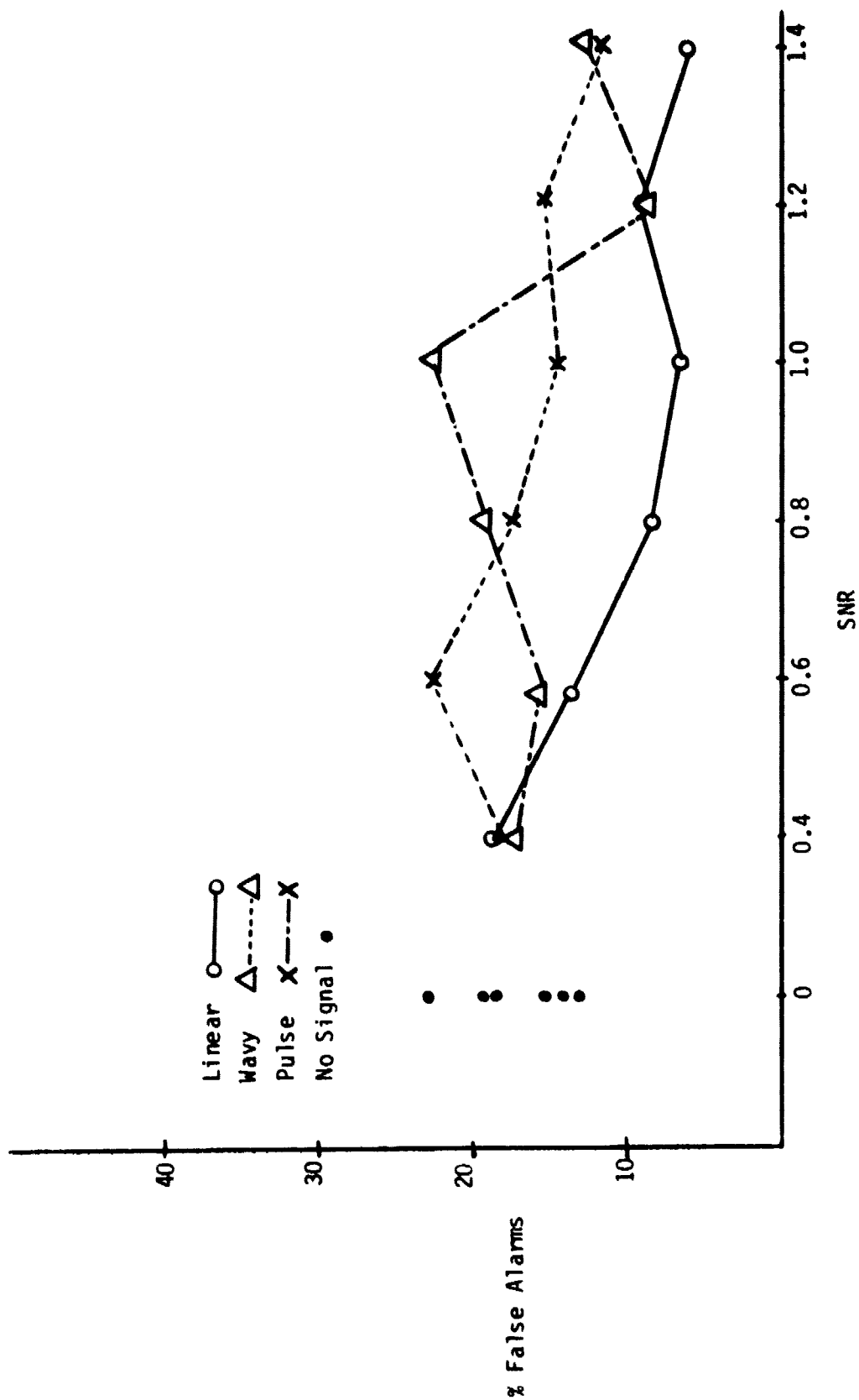


Figure A5.1.6

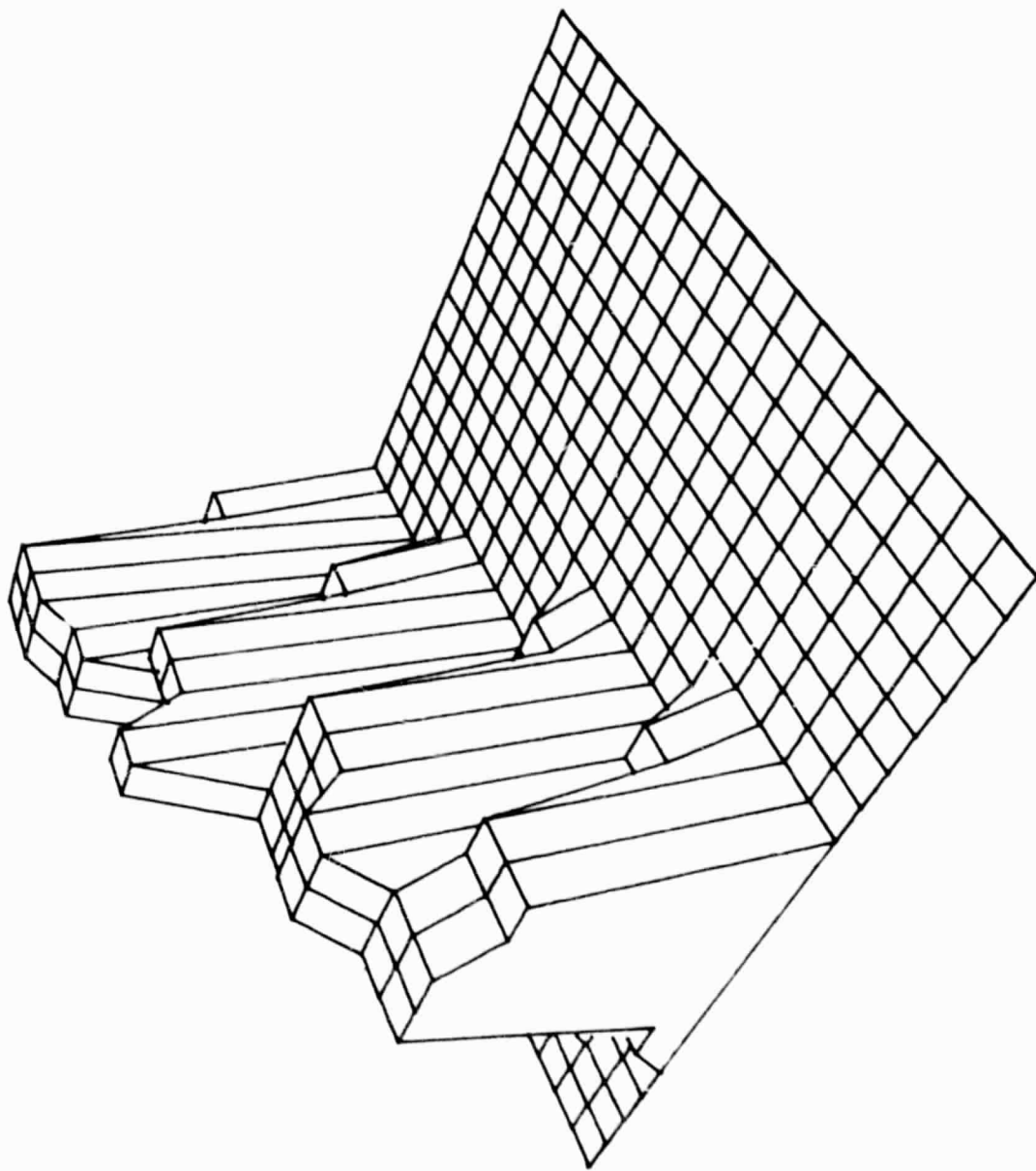


Figure A5.1.7

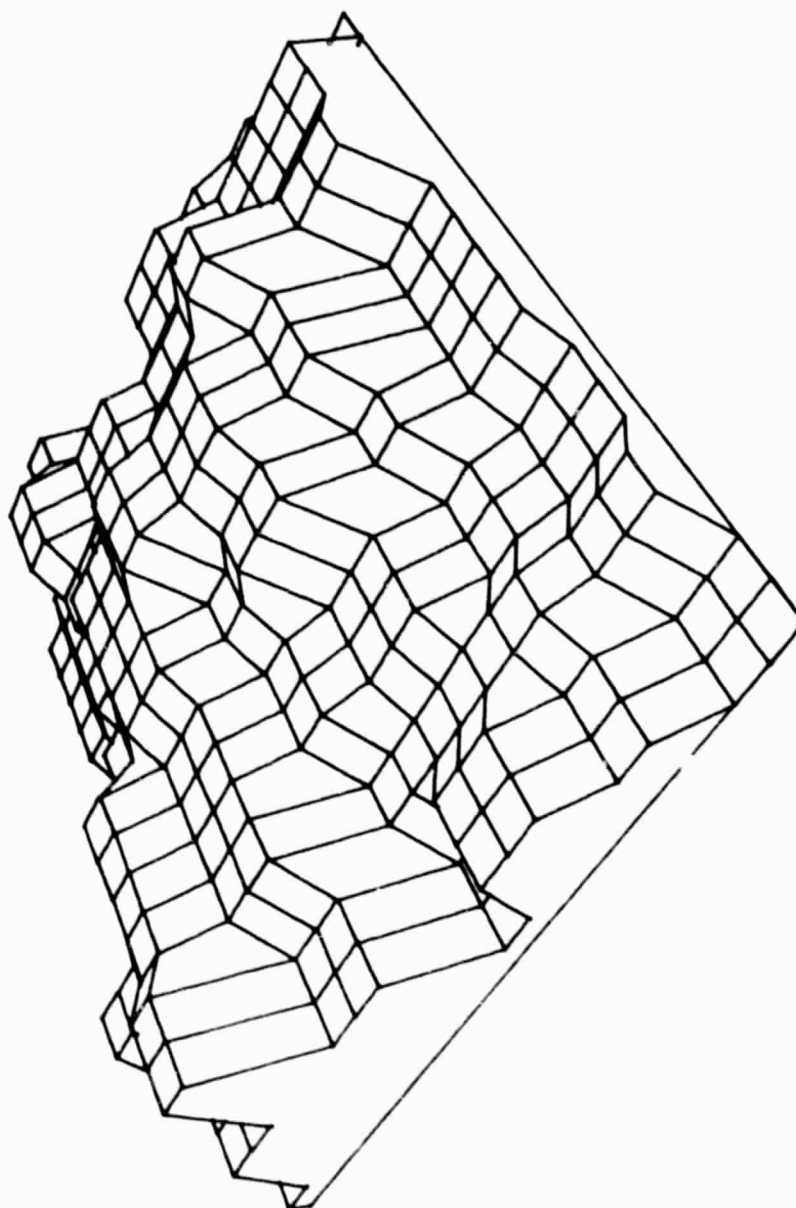


Figure A5.1.1.8

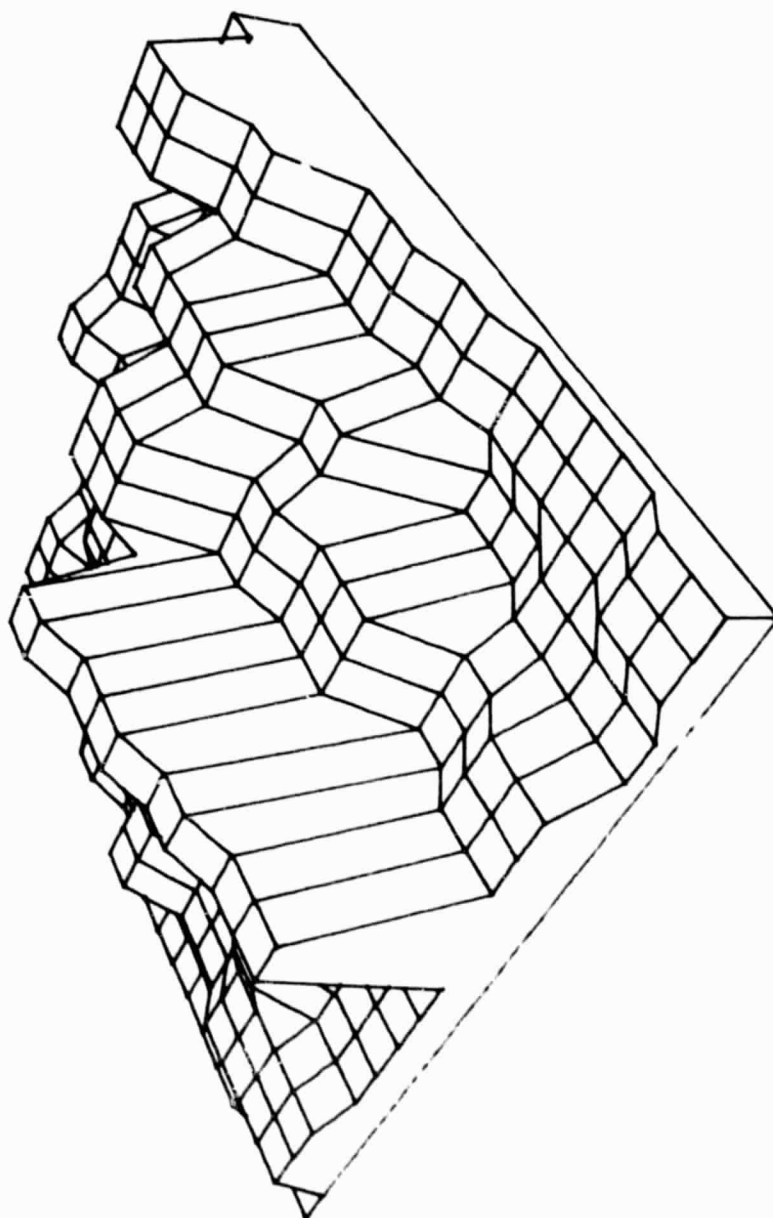


Figure A5.1.9



## Appendix A7.1

### Experimental Procedure Outlines for Several Feasibility Studies of Non-Astrophysical Uses of the Multichannel Spectrum Analyzer and Signal Detection System

Two experimental outlines for preliminary feasibility studies of non-astrophysical uses of the multichannel spectrum analyzer and signal detection system are presented in this appendix. Each of these studies can be implemented with off-the-shelf hardware, however the studies will provide insight into the nature and depth of the results that could be obtained by implementing the full system as proposed in this study.

#### A Source Mechanism Study

In this study, the apparent downward frequency shift which accompanies growth of the acoustic emission source will be investigated.

Parallelepipeds or right circular cylinders will be cut from a single crystal of quartz ( $\text{SiO}_2$ ). A stress riser hole will be drilled perpendicular to the long axis of the sample, which will allow a crack to stably propagate from the stress riser hole when the sample is loaded. Zirconium titanate piezoelectric transducers will be bonded to the sides of the sample with conductive epoxy. The transducer outputs will be amplified by an 80-110 dB preamplifier and recorded on an analog instrument tape recorder. A video camera will be used to monitor crack velocity. The time gate signal on the tape recorder will be used as an input to the camera in order to insure synchronization of the camera and recorder.

Each sample will be loaded perpendicular to the stress rider hole and the long axis of the sample until the crack propagation stress is reached.

When crack propagation is initiated, the sample will be statistically loaded. All acoustic emissions during crack propagation will be recorded on magnetic tape. The power spectrum as a function of crack length will be analyzed with a spectrum analyzer. An off-the-shelf analyzer can be used for the initial study because while the frequency resolution over the entire spectrum of emissions is poor, high frequency resolution can be obtained if the bandwidth of the analyzer is reduced. The spectrum analyzer can be used in narrowband mode to examine the high and low ends of the spectrum by replaying the tape into the spectrum analyzer, in order to determine the relationship between crack length and frequency shift or crack velocity and frequency shift. The sample and experimental configuration are shown in Figure 7.1.1.

#### Stress Field Measurement

Stress field measurement using the Kaiser effect as a stress measurement tool will be investigated. The Kaiser effect is basically an acoustic memory; if a rock has been loaded to 70% of its compressive strength (with acoustic emission activity from roughly 40 - 50% of the compressive strength up to failure) acoustic emission activity will not occur, upon reloading, until the previous load is exceeded.

Parallelepipeds or right circular cylinders will be cut from a relatively homogeneous rock such as sandstone. Each sample will be loaded axially to 70% of the compressive strength of the material and then unloaded. A piezo-electric transducer will be bonded to the side of the sample with conductive epoxy. Each sample will then be reloaded along the same axis up to 85% of the compressive strength of the rock, and the acoustic emissions during reloading

will be monitored. A series of sample sets with varying durations between initial loading and reloading can be run in order to determine the nature and longevity of the Kaiser effect. Samples can also be axially deformed under confining pressure in order to determine if the three dimensional stress field can be evaluated with the Kaiser effect. The output of the transducer will be amplified by an 80-120 dB pre-amplifier and recorded on an analog instrument tape recorder. The signal gating track on the tape recorder will be used to enter the stress level data and time signal. The tape will be replayed through a spectrum analyzer. An off-the-shelf spectrum analyzer will be used. The gross characteristics of the spectrum can be determined with the spectrum analyzer in wideband mode and then the spectrum can be re-examined section by section with the spectrum analyzer in the narrow-band mode. The multichannel spectrum analyzer and signal detector proposed in this study has far greater sensitivity and frequency resolution than an off-the-shelf spectrum analyzer, however the feasibility of using the system to determine stress levels in fault zones can be investigated with off-the-shelf equipment. The sample and experimented configuration are shown in Figure 7.1.2.

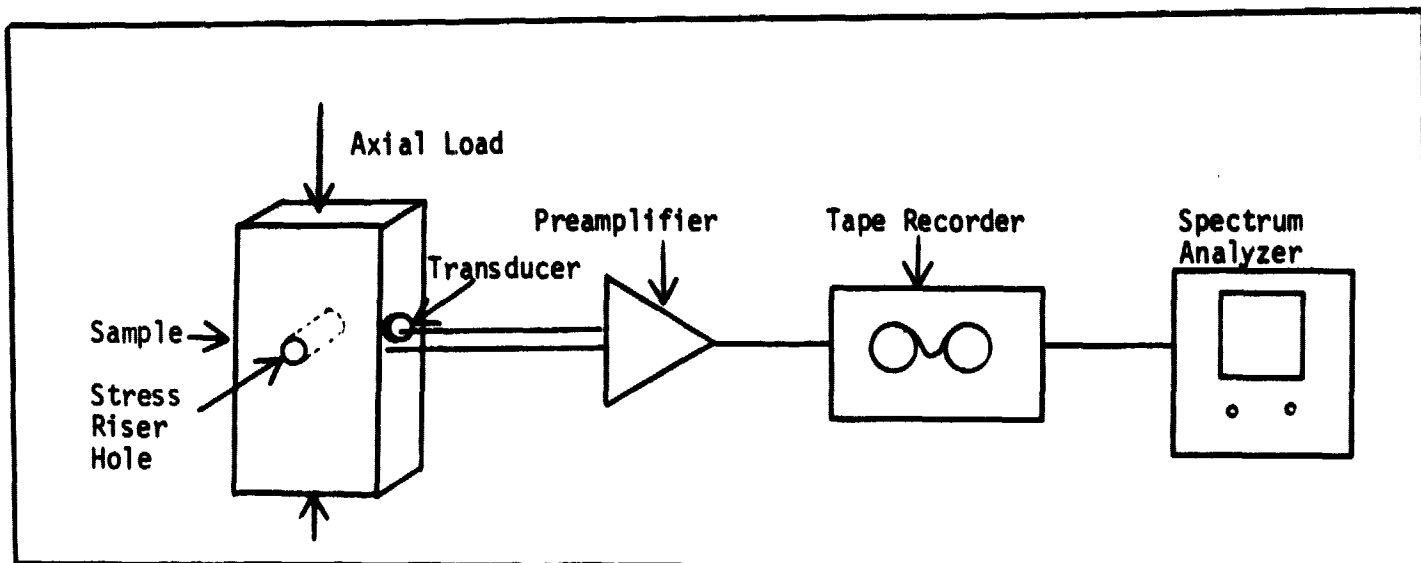


Figure 7.1.1

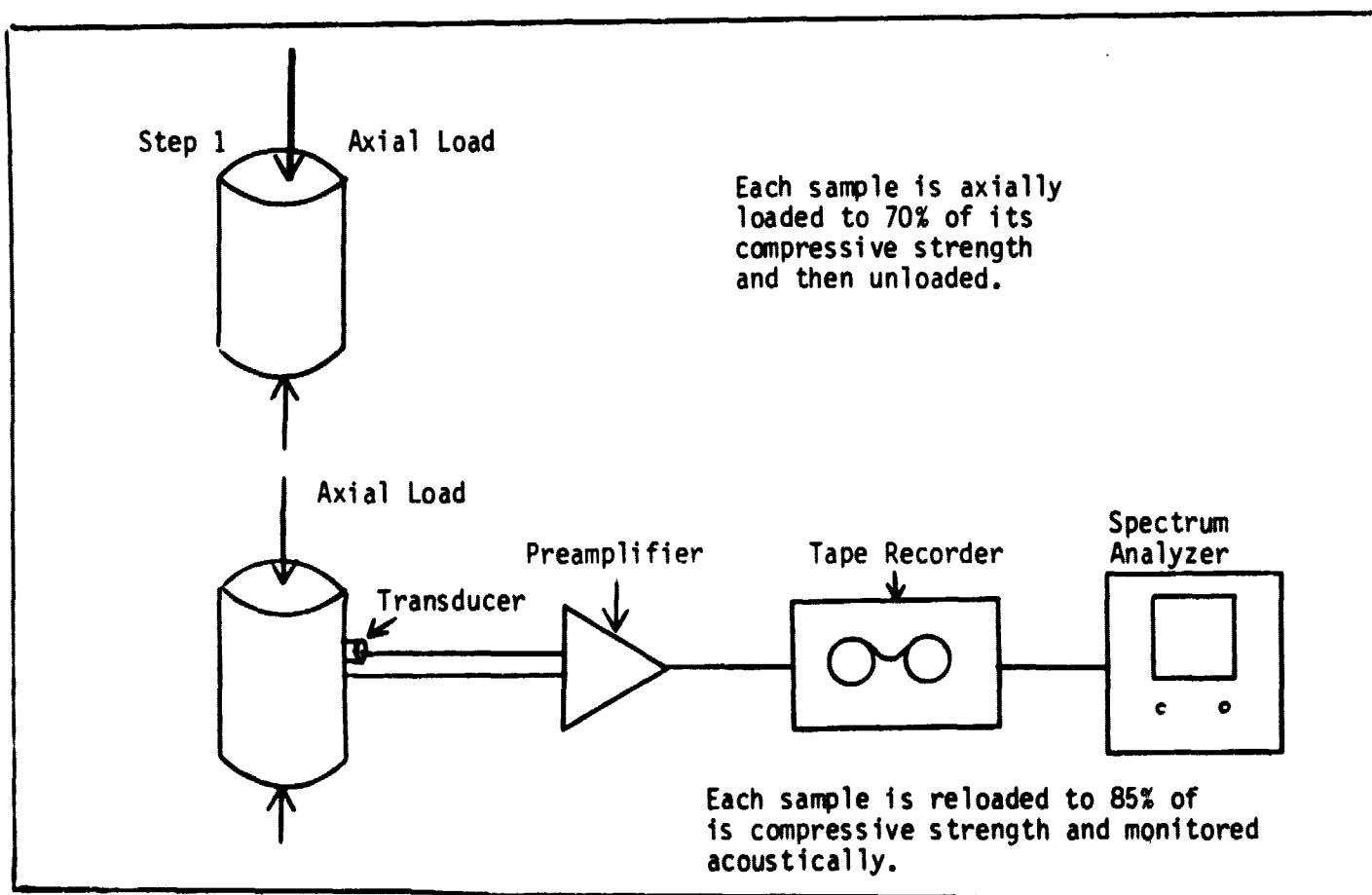


Figure 7.1.2

## Index

- alarm 15, 33, 48, 149, 161, 187, 202
- alternation of search strategies 194
- Analysis of Variance (ANOVA) 67+,  
93, 297+, 302+
  - broadband signals 68, 73
  - Column Effect 69, 94, 161
  - F-ratio 93, 299, 303
  - Interaction Effect 69, 94, 161
  - monochromatic signals 68
  - narrowband signals 73
  - performance 72
  - rapidly drifting signals 68, 73
  - Row Effect 69, 94, 161
  - variance ratio 303
- angular modulation 35
- archive system 48, 168+
- array transposer 121+
- autocorrelation function 16, 61, 288
  
- bandpass signal 258
- baseline removal 74
- beacon, optimal interstellar 8
- block 47, 81, 82, 98
- brightness 191
- broadband signal detector 22, 392+  
(see Pulse Detector)
  
- Carrier Wave Detector (CWD) 43, 47, 77+  
79, 129+
- cell 98, 165, 166
- cluster 100, 167
- cluster analysis 98, 366+
- Cluster Detector (CD) 42, 47, 74,  
162, 164, 361+
- coarse binning 41, 76
- coherence matrix 279
- coherent power 59, 60, 286
- Common Frame of Semantic Reference 10
- Common Processor 149, 151
- communication, one-way 6, 8
- compacted array 98, 361
- Cosmic Search, scientific journal 14
- critical flicker frequency 192
- cross-correlation function 61, 288
  
- data addition 41
- data aggregation 41, 113
- data transfer rate 125
- digital hardware 110
- display parameters 200
  
- display resolution 191
- Doppler shift 357
- double buffering 113, 124
  
- eavesdropping mode 38
- electromagnetic radiation 5
- entire array 98
- Extraterrestrial Intelligence (ETI) 27,40
  
- feedback 193
- filter to detect periodicity in  
signal 33
- Fourier Transform 34, 37, 257, 374
  
- Gaussian distribution of noise 32, 347
- generalized coherence 47, 50+, 61+, 92,  
129+ 131+, 287+, 291+
- Generalized Coherence Value (GCV) 359,  
392+
- geophysical applications
  - acoustic emission analysis 233
  - acoustic events 237
  - acoustic non-destructive testing  
methods 233
  - acoustic/microseismic monitoring  
232
  - burst/event thresholding 236
  - crack propagation 232
  - deformation studies 247
  - dislocation motion 232
  - failure mechanism studies 241
  - felicity studies 243
  - fiber-reinforced plastics 234
  - fracture density studies 243
  - industrial research 235
  - Kaiser effect 234, 245, 424
  - Kaiser studies 242
  - leak detection 242
  - low SNR signal analysis 239
  - microfracturing
  - prediction of catastrophic  
failure 233
  - seismic velocity attenuation  
analysis 238, 244
  - signal amplitude 237
  - source mechanism studies 240,  
246, 422
  - spectrum analysis 237, 239, 246
  - stress field analysis 245, 423
  - stress level 237

geophysical applications  
     stress level 237  
     thermoplastic studies 243  
     threshold counting techniques 236  
     transient elastic waves 232, 244  
  
 histogram 85, 102, 116, 162, 361+  
 human interface, design philosophy 195+  
  
 image processor 200+, 398+  
 input and output data, form of 19  
 interference from artificial satellites 18  
 interference from radio/tv stations 18  
 interstellar chemistry 221  
  
 Lingua Cosmica 10  
 luminance 190  
  
 magnitude--human perception of 188  
 matched filters 374  
 memory systems 113+, 381+  
     disc memory 384  
     Random Access Memory (RAM) 121,  
         124, 151, 173, 384  
     Read-Only Memory (ROM) 127  
     tape memory 385  
 modular system 117  
 modulation rate 35  
 molecules 221  
 monochromatic signals 38  
 m-routine 392+  
 Multichannel Spectrum Analyzer (MCSA) 19  
 multiplex 121  
 narrowband filter  
 narrowband filtering 42  
 narrowband signals 33, 39, 16  
 National Academy of Science 2  
 natural (astronomical) signals 27'  
 noise 268  
     power 32  
     sources of 32  
     thermal 32  
 nuclear quadrapole hyperfine splitting 224  
 Numerical Battery of Tests (NBIT) 47, 74, 79+  
     ANOVA 93, 158  
     Complex Amplitude Coherence 91, 158  
     Degree of Polarization 87  
     Generalized Coherence by Row 92  
     Goodness of Fit 94, 150, 154  
     thresholds 81  
     total power 86, 154  
     8-Hz Pulse Detector 92, 154

Oasis Signal Detector (OSD) 8,  
     13, 45, 73, 110  
 organic molecules 3  
 operator interface 48, 169+, 181+  
 operator monitoring 207+  
     fixed mode 208+  
     free mode 215+  
 operator training 183+  
 Overlap System 147, 151  
  
 pattern block 82, 85, 98, 102, 164  
 pattern class 85, 98, 116, 369  
 pattern point 82, 85, 98, 100  
 pattern recognition 39, 95, 96  
 parent stars of life-bearing  
     planets 4  
 perceptual sensitivity 185  
 perceptual symmetry 184  
 performance degradation 182  
 periodicity 33  
 phase correction 335+  
 pipeline 113, 117, 135, 277, 399  
 pixel (picture element) 321, 331  
 Poincare sphere 51, 275  
 polarization 35, 41  
     coherence 50, 56  
     comparison of polz.-handling  
         techniques 54+  
     degree of 57, 58, 87, 279,  
         280+, 392+  
     measuring polz. 5i  
     partial polz. 55  
     synthesis 273+  
 polarized feed antennas 15, 51  
 priority interrupt 117, 170  
 Project Cyclops--1971 14  
 Pulse Detector (PD) 47, 76+,  
     79, 126+  
  
 quadrature sampling 258  
  
 Radio Astronomy System 48, 173+  
 Radio Frequency Interference (RFI)  
     17, 27, 48, 171, 172  
 radio recombination lines 226  
 radio signals 6  
 realtime processing 46, 119+  
 recycling between stars and  
     interstellar space 221  
 refresh memory 398

scan 98, 121, 168  
Science Workshops on Interstellar  
Communications 14  
search projects 11, 17  
search strategy 7  
Second Look 113+  
self-coherence, degree of 59  
signal characteristics 32  
signal decoding 37, 40  
signal detection 40, 46  
signal identification 31, 40  
signal-to-noise ratio (SNR) 41, 275,  
284+, 291  
sinc function 320+  
sinusoidal signals 33, 38, 60, 259  
  
threshold 32, 170, 186+ 364  
time coherence 50, 54, 59  
time lag 59  
time shift 290, 289  
  
velocity of molecule 222  
  
Walsh transform 376+  
wideband pulses 15  
  
Zeeman splitting in OH 224